

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 09:35:27 ; Search time 60 seconds  
(without alignments)  
3475.334 Million cell updates/sec

Title: US-10-049-957-4

Perfect score: 3936

Sequence: 1 MRGPGCALWLLALRTVLGG.....APLLPLLPALAAARLLPPAL 738

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003as.\*
- 7: Geneseq2003bs.\*
- 8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Query % | Score | Match | Length | ID       | Description |
|------------|---------|-------|-------|--------|----------|-------------|
| 1          | 3936    | 100.0 | 738   | 2      | AAR47899 | Human mel   |
| 2          | 3936    | 100.0 | 738   | 4      | AAB62881 | Membrane    |
| 3          | 3936    | 100.0 | 738   | 5      | AAU78363 | Cell diff   |
| 4          | 3918    | 99.5  | 738   | 1      | AAU70382 | Sequence    |
| 5          | 3834    | 97.4  | 718   | 4      | AAU73169 | Human MTF   |
| 6          | 3441    | 87.4  | 736   | 4      | AAB62880 | Membrane    |
| 7          | 3441    | 87.4  | 736   | 5      | AAU78362 | Cell diff   |
| 8          | 3300    | 83.8  | 738   | 4      | AAB62882 | Membrane    |
| 9          | 3300    | 83.8  | 738   | 4      | AAE06668 | Mouse p97   |
| 10         | 1850    | 47.0  | 502   | 2      | AAU43605 | Amino aci   |
| 11         | 1850    | 47.0  | 502   | 2      | AAU97616 | Amino ter   |
| 12         | 1850    | 47.0  | 502   | 2      | AAU97611 | Amino ter   |
| 13         | 1850    | 47.0  | 502   | 2      | AAU41131 | CTLA4/p97   |
| 14         | 1850    | 47.0  | 502   | 2      | AAU81585 | CTLA4/p97   |
| 15         | 1850    | 47.0  | 502   | 2      | AAU87561 | CTLA4-p97   |
| 16         | 1295    | 32.9  | 708   | 2      | AAU71037 | Draculin.   |
| 17         | 1271.5  | 32.3  | 698   | 2      | AAU12499 | Human tra   |
| 18         | 1269.5  | 32.3  | 698   | 2      | AAU66492 | Human tra   |
| 19         | 1269.5  | 32.3  | 698   | 2      | AAU4354  | Serotrans   |
| 20         | 1269.5  | 32.3  | 698   | 3      | AAU50717 | Human ser   |
| 21         | 1269.5  | 32.3  | 698   | 6      | ABP72819 | Human tra   |
| 22         | 1269.5  | 32.3  | 698   | 7      | ADD45282 | Human Pro   |
| 23         | 1264.5  | 32.1  | 1410  | 2      | AAU07622 | LDLR/TF c   |
| 24         | 1264.5  | 32.1  | 1410  | 6      | ABU04139 | Human exp   |
| 25         | 1263.5  | 32.1  | 679   | 6      | ABP72820 | Human tra   |

|    |        |      |      |   |          |           |
|----|--------|------|------|---|----------|-----------|
| 26 | 1262.5 | 32.1 | 1074 | 2 | AAU07621 | LDLR/TF c |
| 27 | 1261   | 32.0 | 705  | 2 | AAR22423 | Human lac |
| 28 | 1261   | 32.0 | 705  | 2 | AAU31152 | Human lac |
| 29 | 1260.5 | 32.0 | 708  | 2 | AAR22424 | Human lac |
| 30 | 1260.5 | 32.0 | 708  | 2 | AAU31153 | Human lac |
| 31 | 1257.5 | 31.9 | 711  | 2 | AAR08033 | Human lac |
| 32 | 1255.5 | 31.9 | 711  | 2 | AAR43653 | Lactoferr |
| 33 | 1255.5 | 31.9 | 711  | 2 | AAU09342 | Human lac |
| 34 | 1255.5 | 31.9 | 711  | 2 | AAU57317 | Human lac |
| 35 | 1255.5 | 31.9 | 711  | 2 | AAU86021 | Human lac |
| 36 | 1255.5 | 31.9 | 711  | 3 | AAU08182 | Amino aci |
| 37 | 1255.5 | 31.9 | 711  | 4 | AAU02341 | Human lac |
| 38 | 1254   | 31.9 | 690  | 5 | ABG0724  | Codon opt |
| 39 | 1254   | 31.9 | 690  | 5 | AAE27884 | Human cod |
| 40 | 1252   | 31.8 | 687  | 4 | AAU77911 | Human lac |
| 41 | 1252   | 31.8 | 688  | 4 | AAU77910 | Human lac |
| 42 | 1252   | 31.8 | 689  | 4 | AAU77909 | Human lac |
| 43 | 1252   | 31.8 | 690  | 4 | AAU77908 | Human lac |
| 44 | 1252   | 31.8 | 692  | 3 | AAU58733 | Human lac |
| 45 | 1252   | 31.8 | 692  | 4 | AAU97382 | Human lac |

## ALIGNMENTS

RESULT 1  
AAR47899 standard; protein; 738 AA.

AC AAR47899;

DT 25-MAR-2003 (revised)

DT 18-AUG-1994 (revised)

DT 20-JUL-1994 (first entry)

XX Human melanoma-associated antigen p97.

DE Human melanoma-associated antigen p97;

KW Human melanoma-associated antigen p97; Melanotransferrin; GPI;

KW glycosyl-phosphatidylinositol anchor; truncated; soluble;

KW blood-brain barrier; iron binding; Alzheimer's disease; iron metabolism;

KW disorder; haemochromatosis; ischaemic tissue damage; heart disease;

XX skin cancer; brain tumour.

OS Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..19

FT Protein 20..738

FT /label= p97

XX WO9401463-A1.

PD 20-JAN-1994.

PF 09-JUL-1993; 93WO-CA0000272.

XX 10-JUL-1992; 92US-00912291.

XX (UYBR-) UNIV BRITISH COLUMBIA.

PA Jefferies WA, McGeer PL, Rothenberger S, Food MR, Yamada T;

PI WPI: 1994-034993/04.

XX N-PSDB; AAQ55768.

DR New GPI-anchored and soluble forms of p97 - for treating disorders of

XX iron metabolism, delivering drugs to the brain and for diagnosis and

PT treatment of Alzheimer's disease.

XX Disclosure; Page 102-106; 166pp; English.

PS The melanoma-associated antigen p97 has been found to be a GPI-anchored

XX protein expressed on the cell surface which is able to bind iron. A

CC

CC soluble form of p97 can be isolated from the aqueous phase after Triton-X  
 CC -114 phase separation. The soluble (hydrophilic) form of p97 does not  
 CC contain ethanolamine and it has a slower rate of transport than GPI-  
 CC anchored p97. The role of p97 in iron-transport suggests a use in  
 CC modulating iron uptake by cells; p97, its agonists, antagonists and  
 CC stimulants may be useful in treatments where iron-metabolism  
 CC is disturbed, e.g. haemochromatosis and skin cancer. Expression of p97  
 CC (and transferrin receptors) by reactive microglial cells associated with  
 CC senile plaques in Alzheimer's disease (AD) provides a means of diagnosing  
 CC AD. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 738 AA;

Query Match 100.0%; Score 3936; DB 2; Length 738;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGPGALWLLALRTVLGGMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTS 60  
 DB 1 MRGPGALWLLALRTVLGGMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTS 60

QY 61 DHCVOLIAAQEADAITLDGGAIEYAGKEHGLKPVVGEVYDQEVGTSYYAVAVRRSSHT 120  
 DB 61 DHCVOLIAAQEADAITLDGGAIEYAGKEHGLKPVVGEVYDQEVGTSYYAVAVRRSSHT 120

QY 121 IDTLKGVSCHTGINRTVGMNVPVGYLVSRLSVMGCDVLKAVSDYFGGSCVPAGETS 180  
 DB 121 IDTLKGVSCHTGINRTVGMNVPVGYLVSRLSVMGCDVLKAVSDYFGGSCVPAGETS 180

QY 181 YSESLCRLCRGDSGEGVCDKSPLEYYDYSGAFCLAEAGADVAFVHSTVLENTDGT 240  
 DB 181 YSESLCRLCRGDSGEGVCDKSPLEYYDYSGAFCLAEAGADVAFVHSTVLENTDGT 240

QY 241 LPSWGQALLSQDFELLCCRDGSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLNEQ 300  
 DB 241 LPSWGQALLSQDFELLCCRDGSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLNEQ 300

QY 301 RLFSHEGSSQFMSSEAYGQKDLFPKSTSELVPIATQTYEAWLGHEYLHANKGLLCDPN 360  
 DB 301 RLFSHEGSSQFMSSEAYGQKDLFPKSTSELVPIATQTYEAWLGHEYLHANKGLLCDPN 360

QY 361 RLPPYLRCVLTSTPEIQKCGDMAVAFRRQRLKPEIQCVSAKSPQCHMERIOAEQVDVTL 420  
 DB 361 RLPPYLRCVLTSTPEIQKCGDMAVAFRRQRLKPEIQCVSAKSPQCHMERIOAEQVDVTL 420

QY 421 SGEDIYTAGKYGLVPAAGEHYPEDSSNSYYVAVVRRDSSHAFTLDLGRKRSCHAGF 480  
 DB 421 SGEDIYTAGKYGLVPAAGEHYPEDSSNSYYVAVVRRDSSHAFTLDLGRKRSCHAGF 480

QY 481 GSPAGWDVPVGGALIORGFIRPKDCDVLTAVERSEFFNASCVPVNNPKNYPSSLCALCVGDEQ 540  
 DB 481 GSPAGWDVPVGGALIORGFIRPKDCDVLTAVERSEFFNASCVPVNNPKNYPSSLCALCVGDEQ 540

QY 541 GRNKCUGNSQERYGVRGAFCLVENAGDVAFRHTTVPDNTNGNSPAAELRSEDYE 600  
 DB 541 GRNKCUGNSQERYGVRGAFCLVENAGDVAFRHTTVPDNTNGNSPAAELRSEDYE 600

QY 601 LLCPNGARAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKQADLFGDDHNNKGFPM 660  
 DB 601 LLCPNGARAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKQADLFGDDHNNKGFPM 660

QY 661 FDSNNYHGQDLFPKATVRAVPVGEKTYRGWLGLDYVAALGEMSSQCSGAAAPAGAP 720  
 DB 661 FDSNNYHGQDLFPKATVRAVPVGEKTYRGWLGLDYVAALGEMSSQCSGAAAPAGAP 720

QY 721 LPLPLLALAAALLPPAL 738  
 DB 721 LPLPLLALAAALLPPAL 738

RESULT 2  
 AAB62881  
 ID AAB62881 standard; protein; 738 AA.

XX AAB62881;  
 XX 21-MAY-2001 (first entry)  
 XX Membrane bound transferrin like protein amino acid sequence.  
 XX Chondrogenesis promoter; membrane-bound transferrin-like protein; MTF;  
 KW Chondrogenesis regulator; Mtf activator; bone metabolism; human;  
 KW chondral differentiation inhibitor; bone disease.  
 XX Homo sapiens.  
 XX WO2000113951-A1.  
 XX 01-MAR-2001.  
 XX 21-AUG-2000; 2000WO-JP005590.  
 XX 19-AUG-1999; 99JP-00232966.  
 XX (CHUS ) CHUGAI SEIYAKU KK.  
 XX Kato Y, Fujimoto K;  
 XX WPI; 2001-218409/22.  
 XX N-PSDB; AAF62197.  
 XX Chondrogenesis promoters containing membrane-bound transferrin-like  
 PT protein, useful in diagnosis, prevention and treatment of diseases due to  
 PT abnormal chondral metabolism and bone metabolism.  
 XX Claim 3; Page 37-41; 57pp; Japanese.  
 XX This invention relates to chondrogenesis promoters containing a membrane-  
 CC bound transferrin-like protein (MTF). Chondrogenesis promoters,  
 CC chondrogenesis regulators, Mtf activators, Mtf antagonist-containing  
 CC chondral differentiation inhibitors are useful in diagnosis, prevention  
 CC and treatment of diseases due to abnormal chondral metabolism and bone  
 CC metabolism e.g. bone diseases. The present sequence represents the amino  
 CC acid sequence of human MTF  
 XX Sequence 738 AA;

Query Match 100.0%; Score 3936; DB 4; Length 738;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGPGALWLLALRTVLGGMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTS 60  
 DB 1 MRGPGALWLLALRTVLGGMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTS 60

QY 61 DHCVOLIAAQEADAITLDGGAIEYAGKEHGLKPVVGEVYDQEVGTSYYAVAVRRSSHT 120  
 DB 61 DHCVOLIAAQEADAITLDGGAIEYAGKEHGLKPVVGEVYDQEVGTSYYAVAVRRSSHT 120

QY 121 IDTLKGVSCHTGINRTVGMNVPVGYLVSRLSVMGCDVLKAVSDYFGGSCVPAGETS 180  
 DB 121 IDTLKGVSCHTGINRTVGMNVPVGYLVSRLSVMGCDVLKAVSDYFGGSCVPAGETS 180

QY 181 YSESLCRLCRGDSGEGVCDKSPLEYYDYSGAFCLAEAGADVAFVHSTVLENTDGT 240  
 DB 181 YSESLCRLCRGDSGEGVCDKSPLEYYDYSGAFCLAEAGADVAFVHSTVLENTDGT 240

QY 241 LPSWGQALLSQDFELLCCRDGSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLNEQ 300  
 DB 241 LPSWGQALLSQDFELLCCRDGSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLNEQ 300

QY 301 RLFSHEGSSQFMSSEAYGQKDLFPKSTSELVPIATQTYEAWLGHEYLHANKGLLCDPN 360  
 DB 301 RLFSHEGSSQFMSSEAYGQKDLFPKSTSELVPIATQTYEAWLGHEYLHANKGLLCDPN 360

QY 361 RLPPYLRCVLTSTPEIQKCGDMAVAFRRQRLKPEIQCVSAKSPQCHMERIOAEQVDVTL 420

Db 361 RLPPYLRLWCVLSTPEIQKCGDMVAFRRQRKLKPEIQCSAKSPQHCHMERIQAEQVDAVTL 420  
QY 421 SGEDIYTAGKKYGLVPAAGHYAPEDSSNSYYVAVVRRDSSHAFTLDELGRKRSCHAGF 480  
Db 421 SGEDIYTAGKKYGLVPAAGHYAPEDSSNSYYVAVVRRDSSHAFTLDELGRKRSCHAGF 480  
QY 481 GSPAGWDVPVGALIQGFIRPKDCDVLTAVERSEFNASCVPVNNPKNYPSSLCALCVGDEQ 540  
Db 481 GSPAGWDVPVGALIQGFIRPKDCDVLTAVERSEFNASCVPVNNPKNYPSSLCALCVGDEQ 540  
QY 541 GRNKCVCNSQERYGYRGAFRCCLVENAGDVAFVRHTTVPDNTNGHNSPWAELRSEDYE 600  
Db 541 GRNKCVCNSQERYGYRGAFRCCLVENAGDVAFVRHTTVPDNTNGHNSPWAELRSEDYE 600  
QY 601 LLCPCNGARAEVSOFAACNLAIQIPPHAVMVRPDTNIFTVYGLDLDKAODLFGDDHKNKGF 660  
Db 601 LLCPCNGARAEVSOFAACNLAIQIPPHAVMVRPDTNIFTVYGLDLDKAODLFGDDHKNKGF 660  
QY 661 FDSNNYHGQDLLPKDATVRAVPVGEKTYRGWGLDLYVAALGEMSSQCSGAAAPAGAP 720  
Db 661 FDSNNYHGQDLLPKDATVRAVPVGEKTYRGWGLDLYVAALGEMSSQCSGAAAPAGAP 720  
QY 721 LLPLLLPALAARLLPPAL 738  
Db 721 LLPLLLPALAARLLPPAL 738

RESULT 3

AAU78363  
ID AAU78363 standard; protein; 738 AA..

XX AC AAU78363;

XX AC AAU78363;

DT 18-JUN-2002 (first entry)

XX Cell differentiation stimulator associated protein #4.  
DE Cartilage cell differentiation stimulator; osteopathic;

XX Membrane-bound transferrin-like protein; Mtf-BP; concanavalin A; ConA;  
KW membrane bound type transferrin-like protein; Mtf; cartilage disorder;  
KW bone metabolism disease; cell differentiation; cell growth;  
KW extracellular matrix related disease; human.

XX Homo sapiens.

OS

XX

XX

PN JP2002020311-A.

XX 23-JAN-2002.

XX 07-JUL-2000; 2000JP-00206566.

XX 07-JUL-2000; 2000JP-00206566.

PR (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

PA WPI; 2002-287405/33.

XX N-PSDB; ABK12569.

XX

XX

XX

XX

PT A cartilage cell differentiation stimulator useful in the diagnosis of  
PT biophylaxis, cell differentiation, cell growth and construction of  
PT extracellular matrix related diseases.

XX Disclosure; Page 14-16; 17pp; Japanese.

PS

XX

CC The invention describes a cartilage cell differentiation stimulator  
CC (containing a membrane-bound transferrin-like protein (Mtf-BP) and a  
CC membrane bound type transferrin-like protein (Mtf)) and an animal-derived  
CC concanavalin-like drug. The cartilage differentiation stimulator can be  
CC used in diagnosis, prevention and treatment of cartilage and bone  
CC metabolism diseases. They can also be used for diagnosing biophylaxis,  
CC cell differentiation, cell growth and construction of extracellular  
CC matrix related diseases. Mtf-BP strongly stimulates differentiation of

CC cartilage cells and exhibits similar action mechanism with that of plant  
CC derived ConA. This is the amino acid sequence of a cartilage cell  
CC differentiation stimulator associated polypeptide described in the  
CC invention  
XX Sequence 738 AA;  
SQ

Query Match 100.0%; Score 3936; DB 5; Length 738;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGPSALWLLALRTVLGCMVEVRWCATSDDPQHKCGNMSEAFREAGIQPSLLCVRGTS 60

Db 1 MRGPSALWLLALRTVLGCMVEVRWCATSDDPQHKCGNMSEAFREAGIQPSLLCVRGTS 60

QY 61 DHCVCQLIAAQEADAITLDGGAIIYEAQKEHGLKPVVGEVYDQEVGTGYVAVVRRSSHVT 120

Db 61 DHCVCQLIAAQEADAITLDGGAIIYEAQKEHGLKPVVGEVYDQEVGTGYVAVVRRSSHVT 120

QY 121 IDTLKGKVSCHTGINRTVGWNPVGYLVBESGRLSVMGCDVLKAVSDYFGSCVPGAGETS 180

Db 121 IDTLKGKVSCHTGINRTVGWNPVGYLVBESGRLSVMGCDVLKAVSDYFGSCVPGAGETS 180

QY 181 YSESLCRLCRGDSGEGVCDKSPLEYYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGT 240

Db 181 YSESLCRLCRGDSGEGVCDKSPLEYYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGT 240

QY 241 LPSWGOALLSQDFELLCRDGRADVTWQCHLARPAPAHAVVVRADTDGGLIFRLNNEG 300

Db 241 LPSWGOALLSQDFELLCRDGRADVTWQCHLARPAPAHAVVVRADTDGGLIFRLNNEG 300

QY 301 RLFSHEGSSQFMSSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHYHLHANKGLLCDPN 360

Db 301 RLFSHEGSSQFMSSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHYHLHANKGLLCDPN 360

QY 361 RLPPYLRLWCVLSTPEIQKCGDMVAFRRQRKLKPEIQCSAKSPQHCHMERIQAEQVDAVTL 420

Db 361 RLPPYLRLWCVLSTPEIQKCGDMVAFRRQRKLKPEIQCSAKSPQHCHMERIQAEQVDAVTL 420

QY 421 SGEDIYTAGKKYGLVPAAGHYAPEDSSNSYYVAVVRRDSSHAFTLDELGRKRSCHAGF 480

Db 421 SGEDIYTAGKKYGLVPAAGHYAPEDSSNSYYVAVVRRDSSHAFTLDELGRKRSCHAGF 480

QY 481 GSPAGWDVPVGALIQGFIRPKDCDVLTAVERSEFNASCVPVNNPKNYPSSLCALCVGDEQ 540

Db 481 GSPAGWDVPVGALIQGFIRPKDCDVLTAVERSEFNASCVPVNNPKNYPSSLCALCVGDEQ 540

QY 541 GRNKCVCNSQERYGYRGAFRCCLVENAGDVAFVRHTTVPDNTNGHNSPWAELRSEDYE 600

Db 541 GRNKCVCNSQERYGYRGAFRCCLVENAGDVAFVRHTTVPDNTNGHNSPWAELRSEDYE 600

QY 601 LLCPCNGARAEVSOFAACNLAIQIPPHAVMVRPDTNIFTVYGLDLDKAODLFGDDHKNKGF 660

Db 601 LLCPCNGARAEVSOFAACNLAIQIPPHAVMVRPDTNIFTVYGLDLDKAODLFGDDHKNKGF 660

QY 661 FDSNNYHGQDLLPKDATVRAVPVGEKTYRGWGLDLYVAALGEMSSQCSGAAAPAGAP 720

Db 661 FDSNNYHGQDLLPKDATVRAVPVGEKTYRGWGLDLYVAALGEMSSQCSGAAAPAGAP 720

QY 721 LLPLLLPALAARLLPPAL 738

Db 721 LLPLLLPALAARLLPPAL 738

RESULT 4

AAU70382

ID AAP70382 standard; protein; 738 AA.

XX AC AAP70382;

XX AC AAP70382;

DT 25-MAR-2003 (revised)

DT 01-JUL-2002 (revised)

DT 19-FEB-1991 (first entry)



Db 1 MEVWCAATSPQEHKCNMSEAFREAGIQPSLCLVGRGTSADHCVQIAQEAADAITLDDG 60  
 Qy 81 AIYEAGKEHGLKPVGEVYDQEVGTGYAVAVVRRSSHVTIDTLKGKSCHTGINTVGV 140  
 Db 61 AIYEAGKEHGLKPVGEVYDQEVGTGYAVAVVRRSSHVTIDTLKGKSCHTGINTVGV 120  
 Qy 141 NVPVGLVSGRLSVNMGCDLVKAVSDYFGGSCVPGAGETSYSSSLCRLCGDSSGEGVCD 200  
 Db 121 NVPVGLVSGRLSVNMGCDLVKAVSDYFGGSCVPGAGETSYSSSLCRLCGDSSGEGVCD 180  
 Qy 201 KSPLERYDYGAFRCLEAGAGDAFVKHSTVLENTDGTLPWGQALLSODPELLCRDG 260  
 Db 181 KSPLERYDYGAFRCLEAGAGDAFVKHSTVLENTDGTLPWGQALLSODPELLCRDG 240  
 Qy 261 SRADVTEWROCHLARVPAHVVVRAADTDGGLIFRLNNEGQRLFSHSGSSQFMSSEAYGQ 320  
 Db 241 SRADVTEWROCHLARVPAHVVVRAADTDGGLIFRLNNEGQRLFSHSGSSQFMSSEAYGQ 300  
 Qy 321 KOLLFKDSTSELVPIATQTYEAWLGHYELHAMKGLLCDPNRLPPYLRCWCVLSTPEIKCG 380  
 Db 301 KOLLFKDSTSELVPIATQTYEAWLGHYELHAMKGLLCDPNRLPPYLRCWCVLSTPEIKCG 360  
 Qy 381 DMVAFRRORLKEIQCVSAKSPHCHMERIQAEQVDAVTLSDGDIYTAGKKGVLVPAAGE 440  
 Db 361 DMVAFRRORLKEIQCVSAKSPHCHMERIQAEQVDAVTLSDGDIYTAGKKGVLVPAAGE 420  
 Qy 441 HYAPESSSYVAVVVRDSSHAFTLDELGRKRSCHAGFGSPAGWDVPVGLIQRGFTIR 500  
 Db 421 HYAPESSSYVAVVVRDSSHAFTLDELGRKRSCHAGFGSPAGWDVPVGLIQRGFTIR 480  
 Qy 501 PKDCDVLTAVFSEFNASCPVNNPKNYPSSLCALCVGDEGRNKCNGNSOERYYYGRGAF 560  
 Db 481 PKDCDVLTAVFSEFNASCPVNNPKNYPSSLCALCVGDEGRNKCNGNSOERYYYGRGAF 540  
 Qy 561 RCLVENAGDAVFRHTTVPDNTNGHNSPWAELRSEDYELLCPNGARAEVSQFAACNLA 620  
 Db 541 RCLVENAGDAVFRHTTVPDNTNGHNSPWAELRSEDYELLCPNGARAEVSQFAACNLA 600  
 Qy 621 QIPPHAVMVRPDNIITVYGLLDKAQDLFGDDHKNKGFKMFDSSNVHGDQLLFPKDATVRA 680  
 Db 601 QIPPHAVMVRPDNIITVYGLLDKAQDLFGDDHKNKGFKMFDSSNVHGDQLLFPKDATVRA 660  
 Qy 681 VPVGKTTYRGWGLDYVAALGMSQCSGAAAPAPGAPLLPLLLPALAARLLPPAL 738  
 Db 661 VPVGKTTYRGWGLDYVAALGMSQCSGAAAPAPGAPLLPLLLPALAARLLPPAL 718

RESULT 6

AAB62880  
 ID AAB62880 standard; protein; 736 AA.  
 AC AAB62880;  
 XX  
 XX  
 DT 21-MAY-2001 (first entry)  
 DE Membrane bound transferrin like protein amino acid sequence.  
 DE Chondrogenesis promoter; membrane-bound transferrin-like protein; Mtf;  
 KW Chondrogenesis regulator; Mtf activator; bone metabolism; rabbit;  
 KW Chondral differentiation inhibitor; bone disease.  
 XX  
 OS Oryctolagus cuniculus.  
 XX  
 XX WO200113951-A1.  
 XX  
 PD 01-MAR-2001.  
 XX  
 XX 21-AUG-2000; 2000WO-JP005590.  
 XX  
 PR 19-AUG-1999; 99JP-00232966.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.

XX Kato Y, Fujimoto K;  
 PI WPI; 2001-218409/22.  
 XX N-PSDB; AAF62196.  
 DR  
 XX Chondrogenesis promoters containing membrane-bound transferrin-like  
 PT protein, useful in diagnosis, prevention and treatment of diseases due to  
 PT abnormal chondral metabolism and bone metabolism.  
 XX  
 PS Disclosure; Page 32-36; 57pp; Japanese.  
 XX  
 CC This invention relates to chondrogenesis promoters containing a membrane-  
 CC bound transferrin-like protein (Mtf). Chondrogenesis promoters,  
 CC chondrogenesis regulators, Mtf activators, Mtf antagonist-containing  
 CC chondral differentiation inhibitors are useful in diagnosis, prevention  
 CC and treatment of diseases due to abnormal chondral metabolism and bone  
 CC metabolism e.g. bone diseases. The present sequence represents the amino  
 CC acid sequence of rabbit Mtf  
 XX  
 SQ Sequence 736 AA;

Query Match 87.4%; Score 3441; DB 4; Length 736;  
 Best Local Similarity 86.7%; Pred. No. 1.8e-302;  
 Matches 640; Conservative 46; Mismatches 50; Indels 2; Gaps 1;

Qy 1 MRGPSALMLLALRTVLGMEVVRWCATSDPEQHKCNMSEAFREAGIQPSLCLVGRGTS 60  
 Db 1 MCRSAAMWIFLALRTALSGVEVRWCTASEPEQKCEDMSQAFREAGLQPALLCVGTSA 60  
 Qy 61 DHCQVLIQAQDAITLDCGAIYEAGKEHGLKPVGEVYDQEVGTGYAVAVVRRSSHVT 120  
 Db 61 DHCQVLIQAQDAITLDCGAIYEAGKEHGLKPVGEVYDQEVGTGYAVAVVRRSSHVT 120  
 Qy 121 IDTLKGKSCHTGINTVGVNVPVGLVSGRLSVNMGCDLVKAVSDYFGGSCVPGAGETS 180  
 Db 121 IDTLKGKSCHTGINTVGVNVPVGLVSGRLSVNMGCDLVKAVSDYFGGSCVPGAGETS 180  
 Qy 181 YSESLCRLCRGSSGEGVCDKSPLERYDYGAFRCLEAGAGDAVFRHTTVPDNTDGT 240  
 Db 181 YSESLCRLCRGSSGEGVCDKSPLERYDYGAFRCLEAGAGDAVFRHTTVPDNTDGT 240  
 Qy 241 LPSWGQALLSODPELLCRDGSRADVTWROCHLARVPAHVVVRAADTDGGLIFRLNNEG 300  
 Db 241 LPSWGQALLSODPELLCRDGSRADVTWROCHLARVPAHVVVRAADTDGGLIFRLNNEG 300  
 Qy 301 RLFSHSGSSQFMSSEAYGQDLIFKOSTSELVPIATQTYEAWLGHYELHAMKGLLCDPN 360  
 Db 301 RLFSHSGSSQFMSSEAYGQDLIFKOSTSELVPIATQTYEAWLGHYELHAMKGLLCDPN 360  
 Qy 361 RLPPYLRCWCVLSTPEIKQCGDMAVAFRRORLKEIQCVSAKSPHCHMERIQAEQVDAVTL 420  
 Db 361 RLPPYLRCWCVLSTPEIKQCGDMAVAFRRORLKEIQCVSAKSPHCHMERIQAEQVDAVTL 420  
 Qy 421 SGEDIYTAGKKGVLVPAAGEHYAPEDSSSYVAVVVRDSSHAFTLDELGRKRSCHAG 480  
 Db 421 SGEDIYTAGKKGVLVPAAGEHYAPEDSSSYVAVVVRDSSHAFTLDELGRKRSCHAG 480  
 Qy 481 GSPAGWDVPVGLIQRGFTIRPKDCDVLTAVFSEFNASCPVNNPKNYPSSLCALCVGDEQ 540  
 Db 481 GSPAGWDVPVGLIQRGFTIRPKDCDVLTAVFSEFNASCPVNNPKNYPSSLCALCVGDEQ 540  
 Qy 541 GRNKCNGNSOERYYYGRGAFRCLEAGAGDAVFRHTTVPDNTNGHNSPWAELRSEDYE 600  
 Db 541 GRNKCNGNSOERYYYGRGAFRCLEAGAGDAVFRHTTVPDNTNGHNSPWAELRSEDYE 600  
 Qy 601 LLCPNGARAEVSQFAACNLAQIPPHAVMVRPDNIITVYGLLDKAQDLFGDDHKNKGFKM 660  
 Db 601 LLCPNGARAEVSQFAACNLAQIPPHAVMVRPDNIITVYGLLDKAQDLFGDDHKNKGFKM 660  
 Qy 661 FDSSNVHGDQLLFPKDATVRAVPVGKTTYRGWGLDYVAALGMSQCSGAAAPAPGAP 720  
 Db 661 FDSSNVHGDQLLFPKDATVRAVPVGKTTYRGWGLDYVAALGMSQCSGAAAPAPGAP 720

|          |   |   |     |     |               |            |              |            |               |           |            |
|----------|---|---|-----|-----|---------------|------------|--------------|------------|---------------|-----------|------------|
| QY       | 721   | LLPLLALPALAARLLPPAL   | 738 | 181 | YSESLCRLCRDGS | SGEGVCDKSP | PLERYDYDSGAF | CLAEAGADVA | FAVKHSTVLENTD | GKT       | 240        |
| DB       | 721   | LLPLL--PLAAGLLSS  | 736 | 181 | YSESLCRLCRDGT | SGEGVCDKSP | PLERYDYDSGAF | CLAEAGADVA | FAVKHSTVLENTD | ORT       | 240        |
| RESULT 7 |   |   |     |     |               |            |              |            |               |           |            |
| ID       | AAU78362  |   |     |     |               |            |              |            |               |           |            |
| XX       | AAU78362  | standard; protein; 736 AA.                                    |     |     |               |            |              |            |               |           |            |
| AC       | AAU78362;   |   |     |     |               |            |              |            |               |           |            |
| XX       | 18-JUN-2002   | (first entry)   |     |     |               |            |              |            |               |           |            |
| DT       | Cell differentiation stimulator associated protein #3.                    |   |     |     |               |            |              |            |               |           |            |
| DE       | Cartilage cell differentiation stimulator; osteopathic;                   |   |     |     |               |            |              |            |               |           |            |
| XX       | Membrane-bound transferrin-like protein; Mtf-BP; concanavalin A; ConA;    |   |     |     |               |            |              |            |               |           |            |
| KW       | membrane bound type transferrin-like protein; Mtf; cartilage disorder;    |   |     |     |               |            |              |            |               |           |            |
| KW       | bone metabolism disease; cell differentiation; cell growth;               |   |     |     |               |            |              |            |               |           |            |
| XX       | extracellular matrix related disease; rabbit.                             |   |     |     |               |            |              |            |               |           |            |
| XX       | Oryctolagus cuniculus.  |   |     |     |               |            |              |            |               |           |            |
| OS       | JP2002020311-A.   |   |     |     |               |            |              |            |               |           |            |
| XX       | 23-JAN-2002.  |   |     |     |               |            |              |            |               |           |            |
| XX       | 07-JUL-2000; 2000JP-00206566.   |   |     |     |               |            |              |            |               |           |            |
| PF       | 07-JUL-2000; 2000JP-00206566.   |   |     |     |               |            |              |            |               |           |            |
| XX       | (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.                                   |   |     |     |               |            |              |            |               |           |            |
| XX       | WPI; 2002-287405/33.  |   |     |     |               |            |              |            |               |           |            |
| DR       | N-PSDB; ABK12568.   |   |     |     |               |            |              |            |               |           |            |
| XX       | A cartilage cell differentiation stimulator useful in the diagnosis of    |   |     |     |               |            |              |            |               |           |            |
| PT       | biophylaxis, cell differentiation, cell growth and construction of        |   |     |     |               |            |              |            |               |           |            |
| PT       | extracellular matrix related diseases.                                    |   |     |     |               |            |              |            |               |           |            |
| XX       | Disclosure; Page 11-13; 17pp; Japanese.                                   |   |     |     |               |            |              |            |               |           |            |
| PS       | The invention describes a cartilage cell differentiation stimulator       |   |     |     |               |            |              |            |               |           |            |
| CC       | (containing a membrane-bound transferrin-like protein (Mtf-BP) and a      |   |     |     |               |            |              |            |               |           |            |
| CC       | membrane bound type transferrin-like protein (Mtf)) and an animal-derived |   |     |     |               |            |              |            |               |           |            |
| CC       | concanavalin-like drug. The cartilage differentiation stimulator can be   |   |     |     |               |            |              |            |               |           |            |
| CC       | used in diagnosis, prevention and treatment of cartilage and bone         |   |     |     |               |            |              |            |               |           |            |
| CC       | metabolism diseases. They can also be used for diagnosing biophylaxis,    |   |     |     |               |            |              |            |               |           |            |
| CC       | cell differentiation, cell growth and construction of extracellular       |   |     |     |               |            |              |            |               |           |            |
| CC       | matrix related diseases. Mtf-BP strongly stimulates differentiation of    |   |     |     |               |            |              |            |               |           |            |
| CC       | cartilage cells and exhibits similar action mechanism with that of plant  |   |     |     |               |            |              |            |               |           |            |
| CC       | derived ConA. This is the amino acid sequence of a cartilage cell         |   |     |     |               |            |              |            |               |           |            |
| CC       | differentiation stimulator associated polypeptide described in the        |   |     |     |               |            |              |            |               |           |            |
| CC       | invention   |   |     |     |               |            |              |            |               |           |            |
| XX       | Sequence 736 AA;  |   |     |     |               |            |              |            |               |           |            |
| SQ       | Query Match   |   |     |     |               |            |              |            |               |           |            |
|          | Best Local Similarity   | 87.4%; Score 3441; DB 5; Length 736;                          |     |     |               |            |              |            |               |           |            |
|          | Matches 640; Conservative 46; Mismatches 50; Indels 2; Gaps 1;            |   |     |     |               |            |              |            |               |           |            |
| QY       | 1   | MKGPSGALWLLALRTVLGGMEVRVCATSDPROHKCGNSENSEAFREAGTQPSLLCVRTGSA | 60  | 241 | LPSWGQALLSOD  | FELLCRDGS  | RADVTEWRO    | CHLARVPA   | HAVVVRADT     | DGGIIFRL  | NEGQ       |
| DB       | 1   | MKCRSAAMWIFLALRTALGVSVEVRCTASEPEQQCEDMSQAFREAGLQPALLCVQGTSA   | 60  | 241 | LPSWGMLMSR    | DFELLCRDGS | SRASVTEW     | QHCHLARV   | PAHVVVRADT    | DAGLIIFRL | NEGQ       |
| QY       | 61  | DHCVOLIAOEAADATLDGAIYVAGKEHGLKPVVGVYDQEVGTYSYAVAVVRSSHTV      | 120 | 301 | RLFSHEGSS     | FQMFSEAY   | GQKDL        | LFKST      | SELVPIATQ     | TYEAWL    | GHEVYHAKGL |
| DB       | 61  | DHCVQLIAAHEADATLDGAIYVAGKEHGLKPVVGVYDQEVGTYSYAVAVVRSSHTV      | 120 | 301 | RLFSHEGSS     | FQMFSEAY   | GQKDL        | LFKST      | SELVPIATQ     | TYEAWL    | GHEVYHAKGL |
| QY       | 121   | IDTLKGKVSCHTGINRTGWNVPVGYLVESGRLSVNGCDVLKAVSDYFGGSCVPGAGETS   | 180 | 361 | RLPPYLRC      | VLSTPEI    | QKCGDMA      | VAFRQR     | LKPEIQ        | CVSAKS    | POQCMERI   |
| DB       | 121   | INTLAGKVSCHTGINRTGWNVPVGYLVDSGRLSVNGCDVLKAVSEYFGGSCVPGAGETR   | 180 | 361 | RLPPYLRC      | VLSTPEI    | QKCGDMA      | VAFRQR     | LKPEIQ        | CVSAKS    | POQCMERI   |

RESULT 8

AA62882  
ID AA62882 standard; protein; 738 AA.

AC AA62882;

DT 21-MAY-2001 (first entry)

XX Membrane bound transferrin like protein amino acid sequence.

DE Chondrogenesis promoter; membrane-bound transferrin-like protein; Mtf;  
XX Chondrogenesis regulator; Mtf activator; bone metabolism; mouse;  
KW chondral differentiation inhibitor; bone disease.

XX Mus sp.

XX WO200113951-A1.

XX 01-MAR-2001.

XX 21-AUG-2000; 2000WO-JP005590.

XX 19-AUG-1999; 99JP-00232966.

XX (CHUS ) CHUGAI SEIYAKU KK.

XX Kato Y, Fujimoto K;

XX WPI; 2001-218409/22.

XX N-PSDB; AAF62198.





CC regulating CTLA4 receptor positive T cell interactions with B7 receptor  
 CC positive B cells. The method comprises contacting the CTLA4-positive T  
 CC cells with monoclonal antibody fragments reactive with CTLA4. This  
 CC inhibits (and therefore regulates) interactions between CTLA4-positive T  
 CC cells and B7 positive B cells. The method may be used for regulating  
 CC CTLA4 receptor positive T cell interactions with B7 receptor positive B  
 CC cells. In this way the immune system of an individual can be manipulated  
 CC (especially suppressed) for the treatment of autoimmune diseases  
 CC (especially lupus erythematosus) and to prevent host-graft and transplant  
 CC rejection

XX  
 SQ Sequence 502 AA;

Query Match 47.0%; Score 1850; DB 2; Length 502;

Best Local Similarity 100.0%; Pred. No. 2.1e-158;

Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVOLIAAQEADAITLDG 79

DB 156 GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVOLIAAQEADAITLDG 215

QY 80 GAIYEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLGVKSCHTGINRTVG 139

DB 216 GAIYEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLGVKSCHTGINRTVG 275

QY 140 MNVPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYSESLCRLCRGSSGEGVC 199

DB 276 MNVPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYSESLCRLCRGSSGEGVC 335

QY 200 DKSPLEYYDYSAGFRCLAEAGADVAFKVHSTVLENTDGTLPWSQALLSQDFELLCD 259

DB 336 DKSPLEYYDYSAGFRCLAEAGADVAFKVHSTVLENTDGTLPWSQALLSQDFELLCD 395

QY 260 GSRADVTWRQCHLARVPAHVVVRADTDGGLIFRLNNEGQRLFSHEGSSFQMFSEAYG 319

DB 396 GSRADVTWRQCHLARVPAHVVVRADTDGGLIFRLNNEGQRLFSHEGSSFQMFSEAYG 455

QY 320 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 366

DB 456 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 502

RESULT 11

AAW97616

ID AAW97616 standard; protein; 502 AA.

XX

AC AAW97616;

XX

DT 14-JUN-1999 (first entry)

XX

DE Amino terminal CTLA4-carboxy terminal p97 fusion protein.

XX

KW CTLA4 receptor; immunoglobulin; Ig; fusion protein; B7 antigen;  
 KW soluble CTLA4; T cell interaction; B7 positive cell; immune disorder;  
 KW immune system disease; graft versus host disease; psoriasis;  
 KW graft transplant rejection; T cell lymphoma; benign lymphocytic angitis;  
 KW autoimmune disease; lupus erythematosus; Grave's disease;  
 KW Addison's disease; Crohn's disease; multiple sclerosis;  
 KW ulcerative colitis; Sjogren's syndrome; mixed connective tissue disease;  
 KW viral proliferation; T cell activation; AIDS; HTLV1.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN US588579-A.

XX

PD 23-MAR-1999.

XX

PF 08-JUL-1997; 97US-00889666.

XX

PR 27-JUN-1991; 91US-00723617.

PR 22-JAN-1993; 93US-00008898.

PR 15-APR-1994; 94US-00228208.

PR

PR

PR

PR

PR

PR

PR

PR

PR 18-JAN-1995; 95US-00375390.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

PA

PI Damle NK, Kiener PA, Brady W, Ledbetter JA, Linsley PS;

XX WPI; 1999-228484/19.

DR

XX Human CTLA4 receptor protein - used to regulate T-cell interactions with

PT B7 positive cells.

XX

PS Disclosure; Fig 36; 75pp; English.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

RESULT 12

AAW97611

ID AAW97611 standard; protein; 502 AA.

XX

AC AAW97611;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC The specification describes a CTLA4 receptor/immunoglobulin (Ig) fusion  
 CC protein which is reactive with the B7 antigen. DNA encoding the CTLA4  
 CC protein can be used in a host vector system for producing  
 CC soluble CTLA4. The CTLA4 fusion protein can be used for regulating T cell  
 CC interactions with B7 positive cells. The CTLA4Ig fusion protein can be  
 CC used for treating immune system diseases mediated by T cell interactions  
 CC with B7 positive cells. The immune system diseases include graft versus  
 CC host disease, psoriasis, immune disorders associated with graft  
 CC transplant rejection, T cell lymphoma, benign lymphocytic angitis, and  
 CC autoimmune diseases such as lupus erythematosus, Grave's disease,  
 CC Addison's disease, Crohn's disease, multiple sclerosis, ulcerative  
 CC colitis, Sjogren's syndrome, and mixed connective tissue disease. The  
 CC fusion protein may also be used to block the proliferation of viruses  
 CC dependent on T cell activation, such as the virus that causes AIDS,  
 CC HTLV1. The present sequence was created in the course of the invention

XX Sequence 502 AA;

Query Match 47.0%; Score 1850; DB 2; Length 502;

Best Local Similarity 100.0%; Pred. No. 2.1e-158;

Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVOLIAAQEADAITLDG 79

DB 156 GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVOLIAAQEADAITLDG 215

QY 80 GAIYEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLGVKSCHTGINRTVG 139

DB 216 GAIYEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLGVKSCHTGINRTVG 275

QY 140 MNVPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYSESLCRLCRGSSGEGVC 199

DB 276 MNVPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYSESLCRLCRGSSGEGVC 335

QY 200 DKSPLEYYDYSAGFRCLAEAGADVAFKVHSTVLENTDGTLPWSQALLSQDFELLCD 259

DB 336 DKSPLEYYDYSAGFRCLAEAGADVAFKVHSTVLENTDGTLPWSQALLSQDFELLCD 395

QY 260 GSRADVTWRQCHLARVPAHVVVRADTDGGLIFRLNNEGQRLFSHEGSSFQMFSEAYG 319

DB 396 GSRADVTWRQCHLARVPAHVVVRADTDGGLIFRLNNEGQRLFSHEGSSFQMFSEAYG 455

QY 320 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 366

DB 456 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 502

RESULT 12

AAW97611

ID AAW97611 standard; protein; 502 AA.

XX

AC AAW97611;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC The specification describes a CTLA4 receptor/immunoglobulin (Ig) fusion  
 CC protein which is reactive with the B7 antigen. DNA encoding the CTLA4  
 CC protein can be used in a host vector system for producing  
 CC soluble CTLA4. The CTLA4 fusion protein can be used for regulating T cell  
 CC interactions with B7 positive cells. The CTLA4Ig fusion protein can be  
 CC used for treating immune system diseases mediated by T cell interactions  
 CC with B7 positive cells. The immune system diseases include graft versus  
 CC host disease, psoriasis, immune disorders associated with graft  
 CC transplant rejection, T cell lymphoma, benign lymphocytic angitis, and  
 CC autoimmune diseases such as lupus erythematosus, Grave's disease,  
 CC Addison's disease, Crohn's disease, multiple sclerosis, ulcerative  
 CC colitis, Sjogren's syndrome, and mixed connective tissue disease. The  
 CC fusion protein may also be used to block the proliferation of viruses  
 CC dependent on T cell activation, such as the virus that causes AIDS,  
 CC HTLV1. The present sequence was created in the course of the invention

XX Sequence 502 AA;

Query Match 47.0%; Score 1850; DB 2; Length 502;

Best Local Similarity 100.0%; Pred. No. 2.1e-158;

Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVOLIAAQEADAITLDG 79

DB 156 GMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVOLIAAQEADAITLDG 215

QY 80 GAIYEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLGVKSCHTGINRTVG 139

DB 216 GAIYEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLGVKSCHTGINRTVG 275

QY 140 MNVPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYSESLCRLCRGSSGEGVC 199

DB 276 MNVPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYSESLCRLCRGSSGEGVC 335

QY 200 DKSPLEYYDYSAGFRCLAEAGADVAFKVHSTVLENTDGTLPWSQALLSQDFELLCD 259

DB 336 DKSPLEYYDYSAGFRCLAEAGADVAFKVHSTVLENTDGTLPWSQALLSQDFELLCD 395

QY 260 GSRADVTWRQCHLARVPAHVVVRADTDGGLIFRLNNEGQRLFSHEGSSFQMFSEAYG 319

DB 396 GSRADVTWRQCHLARVPAHVVVRADTDGGLIFRLNNEGQRLFSHEGSSFQMFSEAYG 455

QY 320 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 366

DB 456 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 502

RESULT 12

AAW97611

ID AAW97611 standard; protein; 502 AA.

XX

AC AAW97611;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC The specification describes a CTLA4 receptor/immunoglobulin (Ig) fusion  
 CC protein which is reactive with the B7 antigen. DNA encoding the CTLA4  
 CC protein can be used in a host vector system for producing  
 CC soluble CTLA4. The CTLA4 fusion protein can be used for regulating T cell  
 CC interactions with B7 positive cells. The CTLA4Ig fusion protein can be  
 CC used for treating immune system diseases mediated by T cell interactions  
 CC with B7 positive cells. The immune system diseases include graft versus  
 CC host disease, psoriasis, immune disorders associated with graft  
 CC transplant rejection, T cell lymphoma, benign lymphocytic angitis, and  
 CC autoimmune diseases such as lupus erythematosus, Grave's disease,  
 CC Addison's disease, Crohn's disease, multiple sclerosis, ulcerative  
 CC colitis, Sjogren's syndrome, and mixed connective tissue disease. The  
 CC fusion protein may also be used to block the proliferation of viruses  
 CC dependent on T cell activation, such as the virus that causes AIDS,  
 CC HTLV1. The present sequence was created in the course of the invention

XX Sequence 502 AA;

KW graft transplant rejection; T cell lymphoma; benign lymphocytic angitis;  
 KW autoimmune disease; lupus erythematosus; Grave's disease;  
 KW Addison's disease; Crohn's disease; multiple sclerosis;  
 KW ulcerative colitis; Sjogren's syndrome; mixed connective tissue disease;  
 KW viral proliferation; T cell activation; AIDS; HTLV1.

XX Synthetic.  
 OS Homo sapiens.  
 XX US5885796-A.  
 XX 23-MAR-1999.  
 XX 05-JUN-1995; 95US-00465078.  
 XX 27-JUN-1991; 91US-00723617.  
 XX 22-JAN-1993; 93US-00008898.  
 XX 15-APR-1994; 94US-00228208.  
 XX 18-JAN-1995; 95US-00375390.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Damle NK, Brady W, Ledbetter JA, Linsley PS;  
 XX WPI; 1999-228535/19.  
 XX CTLA4 receptor protein for use in treatment of immune system diseases.

XX Disclosure; Fig 36; 75pp; English.  
 XX The specification describes a CTLA4 receptor/immunoglobulin (Ig) fusion  
 CC protein which is reactive with the B7 antigen. DNA encoding the CTLA4  
 CC protein can be used in a vector in a host vector system for producing  
 CC soluble CTLA4. The CTLA4 fusion protein can be used for regulating T cell  
 CC interactions with B7 positive cells. The CTLA4 fusion protein can be  
 CC used for treating immune system diseases mediated by T cell interactions  
 CC with B7 positive cells. The immune system diseases include graft versus  
 CC host disease, psoriasis, immune disorders associated with graft  
 CC transplant rejection, T cell lymphoma, benign lymphocytic angitis, and  
 CC autoimmune diseases such as lupus erythematosus, Grave's disease,  
 CC Addison's disease, Crohn's disease, multiple sclerosis, ulcerative  
 CC colitis, Sjogren's syndrome, and mixed connective tissue disease. The  
 CC fusion protein may also be used to block the proliferation of viruses  
 CC dependent on T cell activation, such as the virus that causes AIDS.  
 CC HTLV1. The present sequence was created in the course of the invention

XX Sequence 502 AA;  
 CC Query Match 47.0%; Score 1850; DB 2; Length 502;  
 CC Best Local Similarity 100.0%; Pred. No. 2.1e-158; Indels 0; Gaps 0;  
 CC Matches 347; Conservative 0; Mismatches 0;  
 QY 20 GMEVWCATSDPEQHKCGNMSFAFRAEAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 79  
 DB 156 GMEVWCATSDPEQHKCGNMSFAFRAEAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 215  
 QY 80 GAIYAGKEHGLKPVGEYDVEGTSYAVAVRRSSHTVITDLTKGVKSCHTGINRTVG 139  
 DB 216 GAIYAGKEHGLKPVGEYDVEGTSYAVAVRRSSHTVITDLTKGVKSCHTGINRTVG 275  
 QY 140 WNPVGYLVESGRSLVWVGCDVLKAVSDYFGGSCVPGAGTSSYSELCLRCRDSGEGVC 199  
 DB 276 WNPVGYLVESGRSLVWVGCDVLKAVSDYFGGSCVPGAGTSSYSELCLRCRDSGEGVC 335  
 QY 200 DKSPLEYYDYSGAFCLAEAGDVAFAVKHSTVLENTDGLTPSGQALLSODPELLCRD 259  
 DB 336 DKSPLEYYDYSGAFCLAEAGDVAFAVKHSTVLENTDGLTPSGQALLSODPELLCRD 395  
 QY 260 GSRADVTWRQCHLARPAPAHAVVVRADTDGGLIFRLNEGQRLFSHEGSSFOFSSAYG 319  
 DB 396 GSRADVTWRQCHLARPAPAHAVVVRADTDGGLIFRLNEGQRLFSHEGSSFOFSSAYG 455  
 QY 320 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAKMGLLCDPNRLPPYL 366

DB 456 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAKMGLLCDPNRLPPYL 502

RESULT 13  
 AAY41131  
 ID AAY41131 standard; protein; 502 AA.  
 XX AAY41131;  
 AC AAY41131;  
 XX 24-JAN-2000 (first entry)  
 DT CTLA4/p97 fusion protein.  
 DE Monoclonal antibody; MAb; extracellular domain; CTLA4; B7 antigen;  
 KW T cell interaction; inflammation; autoimmunity; transplantation; GCHD;  
 KW neoplasia; infectious disease; graft versus host disease; psoriasis;  
 KW immune disorder; lymphoma; leukemia; autoimmune disease; arthritis;  
 KW diabetes mellitus; oncostatin M; fusion protein; p97.

XX Synthetic.  
 OS Homo sapiens.  
 XX US5977318-A.  
 XX 02-NOV-1999.  
 XX 07-JUN-1995; 95US-00488062.  
 XX 27-JUN-1991; 91US-00723617.  
 XX 22-JAN-1993; 93US-00008898.  
 XX 15-APR-1994; 94US-00228208.  
 XX 18-JAN-1995; 95US-00375390.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Kiener PA, Brady W, Damle NK, Linsley PS, Ledbetter JA;  
 XX WPI; 1999-619712/53.  
 XX New anti-CTLA4 monoclonal antibodies, used for treating e.g.  
 CC inflammation, autoimmunity, transplant rejection, infectious diseases or  
 CC neoplasia.

XX Example; Fig 36; 74pp; English.  
 CC The invention provides new monoclonal antibodies (MABs) which bind the  
 CC extracellular domain of CTLA4 and prevent the binding of CTLA4 to B7  
 CC antigen. The MABs can be used for regulating T cell interactions with B7  
 CC positive cells. They can also be used for preventing or reversing  
 CC inflammation and for treating autoimmunity, transplantation, infectious  
 CC diseases and neoplasia. They can be used for treating diseases e.g. graft  
 CC versus host disease (GCHD), psoriasis, immune disorders associated with  
 CC graft transplantation rejection, T cell lymphoma, T cell acute  
 CC lymphoblastic leukemia, testicular angiocentric R cell lymphoma, benign  
 CC lymphocytic angitis, autoimmune diseases such as lupus erythematosus,  
 CC Hashimoto's thyroiditis, primary myxedema, Graves disease, pernicious  
 CC anemia, autoimmune atrophic gastritis, Addison's disease, insulin  
 CC dependent diabetes mellitus, Goodpasture's syndrome, myasthenia gravis,  
 CC pemphigus, Crohn's disease, sympathetic ophthalmia, autoimmune uveitis,  
 CC multiple sclerosis, autoimmune hemolytic anemia, primary biliary  
 CC cirrhosis, idiopathic thrombocytopenia, chronic action hepatitis,  
 CC ulcerative colitis, Sjogren's syndrome, rheumatoid arthritis,  
 CC polymyositis, scleroderma, and mixed connective tissue disease. They can  
 CC also be used for detection, diagnosis, prognosis and monitoring of  
 CC diseases. The present sequence represents the CTLA4/p97 fusion protein  
 CC containing an amino terminal CTLA4 domain and a p97 carboxy-terminal  
 CC domain

XX Sequence 502 AA;  
 SQ Query Match 47.0%; Score 1850; DB 2; Length 502;  
 SQ Best Local Similarity 100.0%; Pred. No. 2.1e-158;

```
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 79
DB 156 GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 215
QY 80 GAIYEAGKEHGLKPVVGEVYDQEVGTSYAVAVVRRSSHVTIDTLKGVSCHTGINRTVG 139
DB 216 GAIYEAGKEHGLKPVVGEVYDQEVGTSYAVAVVRRSSHVTIDTLKGVSCHTGINRTVG 275
QY 140 WNPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYESLCRLCRGDSGEGVC 199
DB 276 WNPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYESLCRLCRGDSGEGVC 335
QY 200 DKSPLEYYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGLIFRLNNEGQRLFSHEGSSQFMSSEAYG 259
DB 336 DKSPLEYYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGLIFRLNNEGQRLFSHEGSSQFMSSEAYG 395
QY 260 GSRADVTWRQCHLARVPAHAVVVRADTDGGLIFRLNNEGQRLFSHEGSSQFMSSEAYG 319
DB 396 GSRADVTWRQCHLARVPAHAVVVRADTDGGLIFRLNNEGQRLFSHEGSSQFMSSEAYG 455
QY 320 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 366
DB 456 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 502

RESULT 14
ID AAW81585
AC AAW81585;
XX
XX 05-FEB-1999 (first entry)
XX
XX CTLA4/p97 fusion protein sequence.
XX
XX CTLA4 receptor; CTLA4-Ig; fusion protein; B7 antigen; hinge; CH2; CH3;
KW extracellular domain; human; immunoglobulin; T cell; immune system;
KW autoimmune disease; cancer; viral infection; p97.
XX
XX Homo sapiens.
XX Unidentified.
XX
XX Key Location/Qualifiers
XX Peptide 1..28
XX Protein /note= "leader sequence"
XX Domain 29..154
XX /note= "CTLA4 partial sequence"
XX Protein 29..154
XX /note= "CTLA4 amino terminal domain"
XX Protein 155..502
XX /note= "p97 partial sequence"
XX Domain 155..502
XX /note= "p97 carboxy terminal domain"
XX
XX US5844095-A.
XX
XX 01-DEC-1998.
XX
XX 18-JAN-1995; 95US-00375390.
XX
XX 27-JUN-1991; 91US-00723617.
XX 22-JAN-1993; 93US-00008898.
XX 28-MAY-1993; 93US-00069693.
XX 15-APR-1994; 94US-00228208.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Brady W, Linsley PS, Damle NK, Ledbetter JA;
XX WPI; 1999-044666/04.
XX
```

```
PT Fusion protein of CTLA4 and immunoglobulin fragment - for treating immune
XX system disorders.
PS
XX Example 8; Fig 36; 75pp; English.
XX
XX This represents a CTLA4/p97 fusion protein. The invention provides a
XX CTLA4-Ig fusion protein that binds the B7 antigen and has a first amino
XX acid sequence consisting of the extracellular domain of CTLA4 and a
XX second amino acid sequence consisting of the hinge, CH2 and CH3 regions
XX of a human immunoglobulin molecule. The fusion protein inhibits
XX interaction of T cells with B7-positive cells and may be useful for
XX treating immune system diseases, e.g. autoimmune diseases, cancer or
XX viral infections. The present sequence is an example of such a fusion
XX protein and contains an amino-terminal CTLA4 domain and a p97 carboxy-
XX terminal domain
XX
XX Sequence 502 AA;
XX
XX Query Match 47.0%; Score 1850; DB 2; Length 502;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-158;
XX Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 79
DB 156 GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 215
QY 80 GAIYEAGKEHGLKPVVGEVYDQEVGTSYAVAVVRRSSHVTIDTLKGVSCHTGINRTVG 139
DB 216 GAIYEAGKEHGLKPVVGEVYDQEVGTSYAVAVVRRSSHVTIDTLKGVSCHTGINRTVG 275
QY 140 WNPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYESLCRLCRGDSGEGVC 199
DB 276 WNPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYESLCRLCRGDSGEGVC 335
QY 200 DKSPLEYYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGLIFRLNNEGQRLFSHEGSSQFMSSEAYG 259
DB 336 DKSPLEYYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGLIFRLNNEGQRLFSHEGSSQFMSSEAYG 395
QY 260 GSRADVTWRQCHLARVPAHAVVVRADTDGGLIFRLNNEGQRLFSHEGSSQFMSSEAYG 319
DB 396 GSRADVTWRQCHLARVPAHAVVVRADTDGGLIFRLNNEGQRLFSHEGSSQFMSSEAYG 455
QY 320 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 366
DB 456 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPNRLPPYL 502

RESULT 15
AAW87561
ID AAW87561 standard; protein; 502 AA.
XX
XX AAW87561;
XX
XX 01-MAR-1999 (first entry)
XX
XX CTLA4-p97 fusion protein.
XX
XX CD28; B7; fusion protein; hinge CH2; CH3; human IgC-gamma1; CTLA4;
KW CTLA4 receptor; ligand; regulation; T-cell interaction; B7-positive cell;
KW immune system disease.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..15
XX /note= "signal peptide"
XX Protein 16..155
XX /note= "CTLA4 protein"
XX Protein 155..502
XX /note= "p97 protein"
XX
XX US5851795-A.
```

XX 22-DEC-1998.  
PD XX  
XX 02-JUN-1995; 95US-00459818.  
PF XX  
XX 27-JUN-1991; 91US-00723617.  
PR XX  
XX 22-JAN-1993; 93US-00008898.  
PR XX  
XX 15-APR-1994; 94US-00228208.  
PR XX  
XX 18-JAN-1995; 95US-00375390.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
XX Ledbetter JA, Brady W, Damle NK, Kiener PA, Linsley PS;  
PI XX  
XX WPI; 1999-080402/07.  
DR XX  
XX Soluble CTLA4 protein that binds B7 antigen of activated B cells - and  
PT fusion proteins useful for regulating T-cell interactions with B cells.  
XX  
XX Claim 16; Fig 36; 75pp; English.  
PS  
XX  
XX The present sequence represents a CTLA4-p97 fusion protein. CTLA4 is a  
CC receptor protein having a high degree of homology with CD28. The CTLA4  
CC receptor is identified as a ligand for the B7 antigen. The CTLA4 protein  
CC can also be used to construct a fusion protein of CTLA4 and human  
CC immunoglobulin (Ig)-gamma1. The CTLA4 protein was fused to the hinge CH2  
CC and CH3 regions of human IgG-gamma1. Soluble CTLA4-Ig fusion proteins can  
CC be used to regulate T-cell interactions with B7-positive cells and to  
CC treat immune system diseases mediated by such interactions  
XX  
SQ Sequence 502 AA;

Query Match 47.0%; Score 1850; DB 2; Length 502;  
Best Local Similarity 100.0%; Pred. No. 2.1e-158;  
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 GNEVWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 79  
Db 156 GNEVWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 215  
QY 80 GAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDTLKGVKSCHTGINRTVG 139  
Db 216 GAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDTLKGVKSCHTGINRTVG 275  
QY 140 WNPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVFCAGETSYSESICRLCRGDSSGEGVC 199  
Db 276 WNPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVFCAGETSYSESICRLCRGDSSGEGVC 335  
QY 200 DKSPLERYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGTLPWSGQALLSQDFELLCRD 259  
Db 336 DKSPLERYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGTLPWSGQALLSQDFELLCRD 395  
QY 260 GSRADVTEWROCHLARVPAHAVVVRADTDGGLIFRLNEGQRLFSEHSSFFQMFSEAYG 319  
Db 396 GSRADVTEWROCHLARVPAHAVVVRADTDGGLIFRLNEGQRLFSEHSSFFQMFSEAYG 455  
QY 320 QKDILFKDSTSELVPIATQTYEAWLGHYHLHAKGLLCDPNRLPPYL 366  
Db 456 QKDILFKDSTSELVPIATQTYEAWLGHYHLHAKGLLCDPNRLPPYL 502

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 09:39:47 ; Search time 21 Seconds  
(without alignments)  
3380.445 Million cell updates/sec

Title: US-10-049-957-4  
Perfect score: 3936  
Sequence: 1 MRGPSALMILLALRTVLGG.....APLLPLLALAAARLLPPAL 738

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: piri:\*

2: piri2:\*

3: piri3:\*

4: piri4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID    | Description        |
|------------|--------|-------------|--------|----------|--------------------|
| 1          | 3936   | 100.0       | 738    | 1 TFHUM  | melanotransferrin  |
| 2          | 1272.5 | 32.3        | 690    | 2 T11749 | transferrin - Atla |
| 3          | 1269.5 | 32.3        | 698    | 1 TFHUP  | transferrin precu  |
| 4          | 1263.5 | 32.1        | 706    | 2 S33761 | transferrin precu  |
| 5          | 1260.5 | 32.0        | 711    | 1 TFHUL  | lactotransferrin p |
| 6          | 1258   | 32.0        | 708    | 2 JC2323 | lactoferrin - goat |
| 7          | 1249   | 31.7        | 708    | 1 TFBOL  | lactotransferrin p |
| 8          | 1228   | 31.2        | 694    | 1 TFRBP  | transferrin precu  |
| 9          | 1213.5 | 30.8        | 696    | 1 S01384 | transferrin - pig  |
| 10         | 1212   | 30.8        | 703    | 2 A45543 | lactoferrin precu  |
| 11         | 1202.5 | 30.6        | 704    | 1 T47228 | carbonic anhydrase |
| 12         | 1202.5 | 30.6        | 705    | 1 TFCHE  | ovotransferrin pre |
| 13         | 1192.5 | 30.3        | 707    | 1 A28438 | lactoferrin precu  |
| 14         | 1160   | 29.5        | 695    | 2 S49163 | transferrin precu  |
| 15         | 1128.5 | 28.7        | 717    | 2 S12100 | transferrin precu  |
| 16         | 824    | 20.9        | 726    | 2 A47275 | transferrin - cock |
| 17         | 551.5  | 14.0        | 681    | 1 A36500 | transferrin precu  |
| 18         | 450    | 11.4        | 1274   | 2 T10729 | transferrin-like p |
| 19         | 408    | 10.4        | 215    | 2 A60166 | hemiferrin - bovin |
| 20         | 408    | 10.4        | 216    | 1 A39684 | hemiferrin - rat   |
| 21         | 397.5  | 10.1        | 311    | 2 A28446 | transferrin - mous |
| 22         | 365    | 9.3         | 629    | 2 S68986 | transferrin - fles |
| 23         | 152.5  | 3.9         | 87     | 2 A39426 | transferrin - shee |
| 24         | 145    | 3.7         | 119    | 2 A39426 | saxiphilin - bullf |
| 25         | 121    | 3.1         | 1251   | 2 A57293 | latent transferrin |
| 26         | 121    | 3.1         | 1639   | 1 MMFFB2 | laminin gamma-1 ch |
| 27         | 119    | 3.0         | 770    | 2 T00204 | LDL receptor relat |
| 28         | 114    | 2.9         | 1562   | 2 T17411 | polyketide synthas |
| 29         | 113.5  | 2.9         | 770    | 2 T00203 | LDL receptor-relat |

## ALIGNMENTS

### RESULT 1

TFHUM

melanotransferrin precursor - human  
N;Alternate names: melanoma-associated antigen gp95/p97

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 28-Jan-2000

C;Accession: A23814; A60925

R;Rose, T.M.; Plowman, G.D.; Teplow, D.B.; Dreyer, W.J.; Hellstrom, K.E.; Brown, J.P.

Proc. Natl. Acad. Sci. U.S.A. 83, 1261-1265, 1986

A;Title: Primary structure of the human melanoma-associated antigen p97 (melanotransfer

A;Reference number: A23814; MUID:86149285; PMID:2419904

A;Accession: A23814

A;Molecule type: mRNA

A;Residues: 1-738 <ROS>

A;Cross-references: EMBL:M12154; NID:g189515; PIDN:AA59992.1; PID:g189518

A;Experimental source: melanoma

R;Furukawa, K.S.; Furukawa, K.; Real, F.X.; Old, L.J.; Lloyd, K.O.

J. Exp. Med. 169, 585-590, 1989

A;Title: A unique antigenic epitope of human melanoma is carried on the common melanoma

A;Reference number: A60925; MUID:89094252; PMID:2463331

A;Accession: A60925

A;Molecule type: protein

A;Residues: 20-25,'X',27-28,'X',30 <FUR>

C;Comment: This protein is found predominantly in human melanomas and in certain fetal

C;Comment: Seven disulfide bonds are predicted in each domain.

C;Genetics:

A;Gene: GDB:MFI2

A;Cross-references: GDB:119387; OMIM:155750

A;Map position: 3q28-3q29

C;Superfamily: transferrin; transferrin repeat homology

C;Keywords: blocked carboxyl end; duplication; glycoprotein; lipoprotein; membrane prot

F;1-19/Domain: signal sequence #status predicted <SIG>

F;18-360/Domain: transferrin repeat homology <TRH1>

F;361-709/Domain: melanotransferrin #status predicted <MTF>

F;711-738/Domain: transferrin repeat homology <TRH2>

F;38,135,515/Binding site: carboxyl-terminal propeptide #status predicted <CTP>

F;710/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature for

Query Match 100.0%; Score 3936; DB 1; Length 738;

Best Local Similarity 100.0%; Pred. No. 7,9e-292;

Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGPSALMILLALRTVLGGMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTS 60

DB 1 MRGPSALMILLALRTVLGGMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTS 60

QY 61 DHCVQLIAAQEADAITLDGGAIVEAGKEHGLKPVGVEVDQEVGTGSYYAVAVRRSSHVT 120

DB 61 DHCVQLIAAQEADAITLDGGAIVEAGKEHGLKPVGVEVDQEVGTGSYYAVAVRRSSHVT 120

QY 121 IDTLKGKVSCHTGINRTGNVFPVGYLVESGRLSVNGCDVLKAVSDYFGSGVCPGAGETS 180

mycosubtilin synth  
probable protein d  
hypothetical prote  
protein disulfide-  
fanconi anemia com  
hypothetical prote  
probable nirB prot  
pyruvate ferredoxi  
MEGF6 protein - ra  
hypothetical prote  
hypothetical prote  
DNA polymerase I p  
hypothetical prote  
polyketide synthas  
alanyl-tRNA synthe  
hepatocyte growth

30 108 2.7 5369 2 T44807  
31 107 2.7 515 2 T06262  
32 106 2.7 1031 2 T43458  
33 104.5 2.7 513 2 T05974  
34 102 2.6 660 2 T02755  
35 102 2.6 661 2 T42754  
36 102 2.6 853 2 H70939  
37 102 2.6 1173 2 B97208  
38 101 2.6 1574 2 T13954  
39 100.5 2.6 444 2 E82499  
40 100.5 2.6 1743 2 T26859  
41 100 2.5 913 2 F82958  
42 100 2.5 1694 2 A83512  
43 100 2.5 8563 2 T30226  
44 99 2.5 929 2 A84379  
45 98.5 2.5 655 1 A46688



R;Park, I.; Schaeffer, E.; Sidoli, A.; Baralle, F.E.; Cohen, G.N.; Zakin, M.M.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 3149-3153, 1985  
 A;Title: Organization of the human transferrin gene: direct evidence that it originated  
 A;Reference number: A94044; MUID:85216459; PMID:3858812  
 A;Accession: A94044  
 A;Molecule type: DNA  
 A;Residues: 73-263, 'E', 265-328, 'N', 330-562 <PAR>  
 A;Cross-references: EMBL:Mil1361  
 R;Adrian, G.S.; Korinek, B.W.; Bowman, B.H.; Yang, F.  
 Gene 49, 167-175, 1986  
 A;Title: The human transferrin gene: 5' region contains conserved sequences which match  
 A;Reference number: A29090; MUID:87192006; PMID:3106157  
 A;Accession: A29090  
 A;Molecule type: DNA  
 A;Residues: 1-72;291-300 <ADR>  
 A;Cross-references: EMBL:M15673  
 R;Uzan, G.; Frain, M.; Park, I.; Besmond, C.; Maessen, G.; Trepot, J.S.; Zakin, M.M.; Ka  
 Biochem. Biophys. Res. Commun. 119, 273-281, 1984  
 A;Title: Molecular cloning and sequence analysis of cDNA for human transferrin.  
 A;Reference number: A32739; MUID:84153910; PMID:6322780  
 A;Accession: A32739  
 A;Molecule type: mRNA  
 A;Residues: 422-690 'G', 692-698 <UZA>  
 A;Cross-references: EMBL:M12525; NID:g339468; PIDN:AAA61142.1; PID:g339469  
 R;MacGillivray, R.T.A.; Mendez, E.; Sinha, S.K.; Sutton, M.R.; Linsback-Zins, J.; Brew,  
 Proc. Natl. Acad. Sci. U.S.A. 79, 2504-2508, 1982  
 A;Title: The complete amino acid sequence of human serum transferrin.  
 A;Reference number: A93911; MUID:92222166; PMID:6953407  
 A;Contents: annotation; disulfide bonds  
 R;Hersberger, C.L.; Larson, J.L.; Arnold, B.; Rosteck, P.R.  
 Ann. N. Y. Acad. Sci. 646, 140-154, 1991  
 A;Title: A cloned gene for human transferrin.  
 A;Reference number: I51959; MUID:92231399; PMID:1809186  
 A;Accession: I51959  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-698 <RES>  
 R;Duguid, J.R.; Bohmont, C.W.; Liu, N.G.; Tourtellotte, W.W.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 7260-7264, 1989  
 A;Title: Changes in brain gene expression shared by scrapie and Alzheimer disease.  
 A;Reference number: I48174; MUID:89386721; PMID:2780570  
 A;Accession: I63133  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 636-696 <RE3>  
 A;Cross-references: GB:M26641; NID:g339988; PIDN:AAA61233.1; PID:g339989  
 R;Schaeffer, E.; Lucero, M.A.; Jeltsch, J.M.; Py, M.C.; Levin, M.J.; Chambon, P.; Cohen,  
 Gene 56, 109-116, 1987  
 A;Title: Complete structure of the human transferrin gene. Comparison with analogous chi  
 A;Reference number: I54011; MUID:88056305; PMID:3678832  
 A;Accession: I54011  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-72 <RE3>  
 A;Cross-references: GB:M17611; NID:g339480; PIDN:AAA61147.1; PID:g339485  
 A;Accession: I68160  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 564-698 <RE4>  
 A;Cross-references: GB:M17614; NID:g339483; PIDN:AAA61148.1; PID:g339486  
 C;Comment: Each of the two repetitive domains binds a ferric ion and a bicarbonate anion  
 C;Genetics:  
 A;Gene: GDB:TF  
 A;Cross-references: GDB:120432; OMIM:190000  
 A;Map position: 3q21-3q21  
 A;Intons: 15/1; 72/3; 119/1; 168/1; 212/2; 231/1; 290/3; 350/1; 401/3; 433/1; 444/1; 49  
 C;Function:  
 A;Description: binds iron for delivery into cells  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein; iron transport; metal binding; plasma  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-698/Product: transferrin #status experimental <MAT>

F;20-350/Domain: transferrin repeat homology <TRH1>  
 F;356-686/Domain: transferrin repeat homology <TRH2>  
 F;28-67,38-58,137-213,156-350,177-193,180-196,190-198,246-260,358-615,364-396,374-387,4  
 F;432,630/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 Query Match 32.3%; Score 1269.5; DB 1; Length 698;  
 Best Local Similarity 40.4%; Pred. No. 1.4e-88;  
 Matches 304; Conservative 118; Mismatches 221; Indels 109; Gaps 32;  
 QY 1 MRGPGALML--LLALRTVLGMEVRCATSDPEQHKCGNMSEAFRE-----AGIOPSL 52  
 DB 1 MRLAVGALLVCAVLGLCLAVDPKTVRCVAHSEATK-----OSFRHMKSVTPSDGPSV 56  
 QY 53 LCVRGTSADHCQVLIAAQEAADAITLDGGAIYEAG-KEHGLKPVVGEVY--DQEVGTSYA 109  
 DB 57 ACVKKASYLDCIRAIANAADAVTLDAGLVDYDAYLAPNNLKPVVAEFYSGKEDPQIFYA 116  
 QY 110 VAVRRSSHVTDITLKGVSCHTGNRTVGNVVPVGLVSVESGRLSVMGCD-----VLK 162  
 DB 117 VAVVKXDSGFQNMQLRGKSCHTGLGRSAGWNIPIGLY-----CDLPEPRKPLEK 167  
 QY 163 AVSDYFGSCVPCGAGETSYSESLCRICRGDSSGEGVCDKSLERYDYDYGAPRCLAEAG 222  
 DB 168 AVANFFSGCAPCADGTDFFQ-LCQLCPG-----CGCSTLMQYFGYGAFLCKLDGAG 219  
 QY 223 DVAFVXHSVTLENTDGTLPFSGQALLSOD-FELLCRDGSRADVTWROCHLARVPAHAV 281  
 DB 220 DVAFVXHSITFENLANKA-----DRDQYELLCNDTRKPKVDEYKDCHLAQVPSHTV 270  
 QY 282 VVRADTDGG---LIFRLNEGQRLFSHEGS-SQMFSSSEAYGOKDLFKDSTSELVPIAT 337  
 DB 271 VAR--SMGKEDLIWELNQAOEHFGKDKSKEQLFSS-PHG-KDLLFKDSAHGFLKVP 326  
 QY 338 Q-TYEAWLGHVYLHMKGLL-----CDPNRLPPVLRMCVLSLTPRIQKCGMAVA 385  
 DB 327 RMDAKMYLGYEYTAIRNLRGTCPAPTDECKP-----VKCALSHHRLKCDWSV- 379  
 QY 386 FRQRILKPIQCVSAKQHCMEIRIOAEQVDAVTLTSGEDIYTAGKYGVLVPAAGEVAPE 445  
 DB 380 ----NSVGKIECVSAETTEDCIKINNGEADAMSLDGGFYIAG-KCGLVPVLAENYKS 434  
 QY 446 DSN-----SYVVAVVRDSSHAFTLDELGRKRSCHAGFGSPAGWVPVGVALLIQGFIR 500  
 DB 435 DNCEDTPEAGYFAVAVVKKASD-LTWDNLKGKKSCHTAGVTRTAGWNI PMGLL-----YNK 489  
 QY 501 PKDCDVLTVASSFFENASCVPVANNPNYPSLLCALCVGDEGGRNKCVCNSOERYVYRGAF 560  
 DB 490 INHC-----RFDEFFSEGCAPGSKK-----DSSLCKLCMG--SGLNLCBPNNKEGYTYGTAF 540  
 QY 561 RCLIVENAGDVAFVRHTTVPDNTNGHNSPAAELRSEDEYELLCPNGARAEVSOFAACNLA 620  
 DB 541 RCLVEK-GDVAFVKHQTPQNTGGKPNPDPWAKNLNEKYELLCCLDCTRKPVVEYANCHLA 599  
 QY 621 QTPPHAVVRPDNTNIPVYVGLDKAQDLFGDDHN--KNGFKMPDSSNYHGQDLFLKDATV 678  
 DB 600 RAPNHAIVTRKDKEA-CVHKILRQQOHLFGSNVTDGSGNFCFLRSET---KDLLFRDDTV 655  
 QY 679 RAVPVGEKTYRGWGLDYVAALGHSQQCS 710  
 DB 656 CLAKLHDRNTYBKLYGEEYVKA VGNL--RKCS 685

RESULT 4

S33761  
 transferrin precursor - horse  
 N;Alternate names: growth-promoting factor  
 C;Species: Equus caballus (domestic horse)  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Aug-1999  
 C;Accession: S33761; S02145  
 R;Carpenter, M.A.; Broad, T.E.  
 Biochim. Biophys. Acta 1173, 230-232, 1993  
 A;Title: The cDNA sequence of horse transferrin.  
 A;Reference number: S33761; MUID:93277958; PMID:8504171



R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.; Eur. J. Biochem. 145, 659-666, 1994  
A;Title: Human lactoferrin: amino acid sequence and structural comparisons with other  
A;Reference number: A31000; MUID:85076667; PMID:6510420  
A;Accession: A31000  
A;Molecule type: protein  
A;Residues: 20-140,142-169;171-203,'L',205,'K',207-208,'K',210-385,'Q',387-391,'W',393-4  
A;Note: this is the final paper in a series  
R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norbkov, L.  
Eur. J. Biochem. 241, 303-308, 1996  
A;Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity  
A;Reference number: S74119; MUID:97054624; PMID:8989821  
A;Accession: S74119  
A;Molecule type: protein  
A;Residues: 'G',23-24,'R',26-27,'XX',30-32 <HOU>  
A;Experimental source: neutrophil granulocytes  
C;Genetics:  
A;Gene: GDB:LTF  
A;Cross-references: GDB:119368; OMIM:150210  
A;Map position: 3q21-3q23  
C;Superfamily: transferrin; transferrin repeat homology  
C;Keywords: duplication; glycoprotein; iron binding; milk  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-711/Product: lactotransferrin #status experimental <MAT>  
F;21-356/Domain: transferrin repeat homology <TRH1>  
F;360-699/Domain: transferrin repeat homology <TRH2>  
F;29-65,39-56,135-218,177-193,190-201,251-265,503-697,595-609/Disulfide bonds: #status  
F;157,498/Binding site: carbohydate (Asn) (covalent) #status experimental  
F;368-400,378-391,425-706,447-669,479-554,513-527,524-537,647-652/Disulfide bonds: #stat

Query Match 32.0%; Score 1260.5; DB 1; Length 711;  
Best Local Similarity 39.8%; Pred. No. 7.1e-88;  
Matches 302; Conservative 115; Mismatches 234; Indels 107; Gaps 29;

QY 6 GALMLLLAURTVLGGMEVWCATSPDEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCQV 65  
DB 12 GAGLCLAGR---RRSVQCAVQSEATKCFQWRMRKVR-GPPVSCIKRDSPIQCIQ 67  
QY 66 LIAAQEADAITLDGGAIEAG-KEHGLKPVVGVVY--DOEVGTSYAVAVVRRSSHVTID 122  
DB 68 AIAENRADAVTLDDGFIYEAGLAPYKLRPVAAEVGTERQPRHYAVAVVKKGSFQLN 127  
QY 123 TLKGVKSCHTGINTRTVGMNVVPGYL-----VESGRLSVMGCDVLRKAVSDYFGG 170  
DB 128 ELQGLKSCHTGLRTAGNVPIGLTRPFLNWTGPPPIEA-----AVARFFSA 175  
QY 171 SCVPGAGETSYSLCRLCRGDSGEGVCDKSPLEYYDYGAFRCLEAGAGDVAFKHS 230  
DB 176 SCVPGADKQGF-PNLCLRCAG--TGENKCAFSSQSBPFYSYGAFKCLRDGAGDVAFIRS 232  
QY 231 TVLENTDGTLPWSGOALLSQDFELLCDGSRADVTETWROCHLARVPAHVVVRA-DTDG 289  
DB 233 TVFEDLSDEA-----ERDEVELLCPDNTKRPVDFKDXCHLARVPSHAVVARSVNGKE 284  
QY 290 GLIFRLNIGORLFSHEGS-SFQMFSSAYGOKLLFKDST---SELVP-IAQTVEAWL 344  
DB 285 DAIWNLLROAQEFKDKSPKOLFQSPS-GQKDLFFKDSAIQFSRVPIDRISGLY--L 340  
QY 345 GHEYLHMKCLLCPNRLP---PYLRWCVLSTPEIQKCGDMAVAFRRQRKPEIQCSAK 401  
DB 341 GSGYFTAIQNLKSEEEVARRARVVMCAVGEOLKCNQWS-----GLSEGSVTCSSAS 395  
QY 402 SPOHCMERTQAEQVDVTLTSGEDIYTAGKKYGLVPAAGEHYAPEDSSN-----S 450  
DB 396 TTEDCIALVLKGEADAMSLDGGVYTAG-KCGLVPVLAENYKSSQSDPDNCPDRPVEG 454  
QY 451 YVAVVRRDSSHAFTLDELGRKRSCHAGFGSPAGWDVPVVGALIQRFIRPKDCDVLTA 510  
DB 455 YLAVAVVRR-SDTSLTWSNVKSKSCTAVDRTAGWNIEMGLL----FMQTGSC----KF 505  
QY 511 SEFFNASCVPVNNPKNYPSSLCALCVGDEGRNKCVCNQSERYGYRGAFRCLEVENAGDV 570  
DB 506 DEYFSQSCAPGSDPR---SNLCALCTGDEQGENKCVNERNERYGYGTGAFRCLEVENAGDV 562

QY 571 AFVRHTTVPDNTNGHNSPEWAAELRSEDYELLCPNGARAEVSQFAACNLAIQIPPHAVMR 630  
DB 563 AFVKDVTVLQNTDGNNEAWAKDLADFALLCLDCKRPVTEARSCHLAAEPNHAHVSR 622  
QY 631 PDTNFTVYVGLDQKQADLFGDDHKNKG-----FMFDSNHYGQDOLLFKDQATVRAVPVG 684  
DB 623 MD-KVERLQVLLHQQAQFG----RNGSDCPDKFCLFQSET---KNLLFNNDTECLARLH 674  
QY 685 EKTITRGMGLDVAALEGMSQQSCGAAAPAPGAPLL 722  
DB 675 GKITYEKLGPQYVAGITNL--KKCS-----TSPLL 703

RESULT 6  
JC2323  
lactoferrin - goat  
C;Species: Capra aegagrus hircus (domestic goat)  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
C;Accession: JC2323  
R;Le Provost, F.; Nocart, M.; Guerin, G.; Martin, P.  
Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
A;Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant loci  
A;Reference number: JC2323; MUID:94380047; PMID:8093048  
A;Accession: JC2323  
A;Molecule type: mRNA  
A;Residues: 1-708 cLEP  
C;Superfamily: transferrin; transferrin repeat homology  
C;Keywords: duplication; glycoprotein  
F;359-696/Domain: transferrin repeat homology <TRH2>  
F;252,300,387,495,564/Binding site: carbohydate (Asn) (covalent) #status predicted

Query Match 32.0%; Score 1258; DB 2; Length 708;  
Best Local Similarity 40.8%; Pred. No. 1.1e-87;  
Matches 305; Conservative 110; Mismatches 243; Indels 90; Gaps 30;

QY 6 GALMLLLAURTVLGGMEVWCATSPDEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCQV 65  
DB 12 GAGLCLAGR---KNVRWCAISLPWMSKCYQWRRMRKLG-APSICTVRRTSALECIR 66  
QY 66 LIAAQEADAITLDGGAIEAGKE-HCLKPVVGEVYDOEVG--TSYYAVAVVRRSSHVTID 122  
DB 67 AIAENRADAVTLDDGFIYEAGLAPYKLRPVAAEVGTERQPRHYAVAVVKKGSFQLN 126  
QY 123 TLKGVKSCHTGINTRTVGMNVVPGYL-----VESGRLSVMGCDVLRKAVSDYFGGSCVPGA 176  
DB 127 QLOQKSCHEMLGRAGWNIPIVILRPLSWTESAE-----PLQAVARFSSASCVPVCV 180  
QY 177 GETSYSESICRLCRGDSGEGVCDKSPLEYYDYGAFRCLEAGAGDVAFKHSTVLENT 236  
DB 181 DGKAY-PNLCLQCKG--VGENKCACSSQBPYFGYSGAFKCLQDQAGDVAFKETTVFEN- 236  
QY 237 DKTLPWSGOALLSQDFELLCDGSRADVTETWROCHLARVPAHVVVRA-DTDGGLIFRL 295  
DB 237 ----LPEKAD---RDQYELLCLNTRAPVDAFKECHLAQVPSHAVVARSVNGKENLIWEL 289  
QY 296 LNEGQRLF-SHEGSSQFQMFSSAYGOKLLFKDSTSELVPIATQTYEA-WLGEHYLHAMK 353  
DB 290 LRKAQEKFGKNSQRFQFSGPE-GRDILLFKDSALGFVRIPSKYDSALYLSRYLTALK 348  
QY 354 GLLCDPNRLP---PYLRWCVLSTPEIQKCGDMAVAFRRQRKPEIQCSAKSPQHMERI 410  
DB 349 NLRTAEELKARCTRVWCAVGPESQSKQWS-----EQSGONVTCATSTTDDCIALV 403  
QY 411 QAEQVDVTLTSGEDIYTAGKKYGLVPAAGEH-----YAPED-----SSNYYVAVVRRDS 461  
DB 404 LRGEADALSDGGYITAG-KCGLVPVMAENRKSYSLSLDCVLRPTEGYLAVAVVVK-A 461  
QY 462 SHAFTLDELGRKRSCHAGFGSPAGWDVPVVGALI-ORGFIRPKDCDVLTVASEFFNASCVP 520  
DB 462 NEGLTWSNLKSKSCTAVDRTAGWNIEMGLTANQTG-----SC-----AFDEFFSOSCAP 512  
QY 521 VNNPKNYPSSLCALCVGDEQGRNKCVCNQSERYGYRGAFRCLEVENAGDVAFVRHTTVPD 580

Db 513 GADPK---SSLCALCAGDQGLDKCVPSKEXYGYTGAFRCLEADVGDAFVKNDTWWE 569  
QY 581 NTNGHNSPWAELRSYDYLCPNGARAEVSQFACNLAQIPPHAVMVRPDTNIFTVYG 640  
Db 570 NTNGESSADWAKNLNREDFRLCLDGTTPVTEAOSCYLAVAPNHAVNSRSDRAAHVEQV 629  
QY 641 LLDKQDLRGDDHNKNG-----EKMFDSNNYHGQDLFPKDATVRAVPVGEKTYRGWLG 694  
Db 630 LHQO-QALFG-----KNGKCPDQFCFLFKSET---KNLFDNDTECLAKLGGRTYKYLK 681  
QY 695 LDYVAALGEMSSQCSGGAAPAGPALL 722  
Db 682 TEYVTAIANL--KKCS-----TSPLL 700  
RESULT 7  
TFEOL  
lactotransferrin precursor - bovine  
N:Alternate names: lactoferrin  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text\_change 11-May-2000  
A:Accession: I45919; S14674; S14110; S18517; JT0595; S13097; S18518; S13881; PL0148; S21  
R:Tsang, T.C.; Burns, D.K.; Wang, F.; Pan, Y.  
FASEB J. 6, 233, 1991  
A:Title: Cloning of a 80-kD advanced glycosylation end product (AGE) binding protein fro  
A:Reference number: I45919  
A:Accession: I45919  
A:Status: translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-708 <TSA>  
A:Cross-references: GB:L08604; NID:gl63269; PIDN:AAA30609.1; PID:gl63270  
R:Piece, A.  
submitted to the EMBL Data Library, November 1990  
A:Reference number: S14674  
A:Accession: S14674  
A:Molecule type: mRNA  
A:Residues: 1-144, 'V', 146-163, 'PP', 166-339, 'A', 341-438, 'Y', 440-513, 'R', 515-708 <PI1>  
A:Cross-references: EMBL:X57084; NID:G505; PIDN:CAA40366.1; PID:G506  
R:Piece, A.; Colavizza, D.; Benaisa, M.; Maes, P.; Tartar, A.; Montreuil, J.; Spik, G.  
Eur. J. Biochem. 196, 177-184, 1991  
A:Title: Molecular cloning and sequence analysis of bovine lactotransferrin.  
A:Reference number: S14110; MUID:91160550; PMID:2001696  
A:Accession: S14110  
A:Molecule type: mRNA  
A:Residues: 3-144, 'V', 146-339, 'A', 341-438, 'Y', 440-513, 'R', 515-708 <PI2>  
A:Cross-references: EMBL:X57084  
A:Accession: S18517  
A:Molecule type: protein  
A:Residues: 20-35, 'DS', 36-38, 'P', 40-86, 'C', 88-708 <ME3>  
R:Goodman, R.E.; Schanbacher, F.L.  
Biochem. Biophys. Res. Commun. 180, 75-84, 1991  
A:Title: Bovine lactoferrin mRNA: sequence, analysis, and expression in the mammary gland  
A:Reference number: JT0595; MUID:92028986; PMID:1718281  
A:Accession: JT0595  
A:Molecule type: mRNA  
A:Residues: 1-65, 'FG', 68-296, 'S', 298-339, 'A', 341-708 <GOO>  
A:Cross-references: GB:IM63502  
N:Note: the authors translated the codon CCG for residue 66 as Arg and TCT for residue 2  
R:Mead, P.E.; Tweedie, J.W.  
Nucleic Acids Res. 18, 7167, 1990  
A:Title: cDNA and protein sequence of bovine lactoferrin.  
A:Reference number: S13097; MUID:91088328; PMID:2263492  
A:Accession: S13097  
A:Molecule type: mRNA  
A:Residues: 28-33, 'DS', 36-38, 'P', 40-708 <MEA>  
A:Cross-references: EMBL:X54801  
A:Accession: S18518  
A:Molecule type: protein  
A:Residues: 20-47; 59-66; 132-139; 256-277; 278, 305-332; 343-351; 361-363; 586, 587-589; 598-619  
R:Mead, P.E.  
submitted to the EMBL Data Library, October 1990  
A:Reference number: S13881

A:Accession: S13881  
A:Molecule type: mRNA  
A:Residues: 28-38, 'P', 40-86, 'C', 88-708 <ME3>  
A:Cross-references: EMBL:X54801  
R:Rejman, J.J.; Hegarty, B.M.; Hurley, W.L.  
Comp. Biochem. Physiol. B 93, 929-934, 1989  
A:Title: Purification and characterization of bovine lactoferrin from secretions of the  
A:Reference number: PL0148; MUID:90031466; PMID:2805645  
A:Accession: PL0148  
A:Molecule type: protein  
A:Residues: 20-27, 'X', 29-37, 'X', 39-54, 'X', 56-59 <REU>  
R:Beilamy, W.; Takase, M.; Yamauchi, K.; Wakabayashi, H.; Kawase, K.; Tomita, M.  
Biochim. Biophys. Acta 1121, 130-136, 1992  
A:Title: Identification of the bactericidal domain of lactoferrin.  
A:Reference number: S21756; MUID:92287941; PMID:1599934  
A:Accession: S21756  
A:Molecule type: protein  
A:Residues: 36-60 <BEL>  
R:Shimazaki, K.; Tanaka, T.; Kon, H.; Oota, K.; Kawaguchi, A.; Maki, Y.; Sato, T.  
J. Dairy Sci. 76, 946-955, 1993  
A:Title: Separation and characterization of the C-terminal half molecule of bovine lacto  
A:Reference number: A56659; MUID:93253156; PMID:8486845  
A:Accession: A56659  
A:Molecule type: protein  
A:Residues: 20-25; 302-308; 359-366, 'X', 368-376, 'X', 378 <SHI>  
C:Superfamily: transferrin; transferrin repeat homology  
C:Keywords: duplication; glycoprotein; iron; iron binding; metalloprotein; milk  
F:11-19/Domain: signal sequence #status predicted <SIG>  
F:20-708/Product: lactotransferrin #status experimental <MAT>  
F:20-355/Domain: transferrin repeat homology <TRH1>  
F:36-60/Region: antimicrobial  
F:359-606/Domain: transferrin repeat homology <TRH2>  
F:28-64; 134-217; 176-192; 179-200; 189-202; 250-264; 367-399, 377-390, 424-703, 444-666, 476-551,  
F:38-55/Disulfide bonds: #status predicted  
F:79, 111, 211, 272/Binding site: iron (Asp, Tyr, Tyr, His) #status experimental  
F:140/Binding site: carbonate (Arg) #status experimental  
F:252, 300, 387, 495, 564/Binding site: carbohydrate (Asn)  
F:414, 452, 545, 614/Binding site: iron (Asp, Tyr, Tyr, His) #status experimental  
F:482/Binding site: carbonate (Arg) #status experimental  
Query Match 31.7%; Score 1249; DB 1; Length 708;  
Best Local Similarity 40.2%; Pred. No. 5.3e-87;  
Matches 300; Conservative 119; Mismatches 241; Indels 86; Gaps 30;  
QY 6 GALMILLALRTVLGMEVNRWCATSDPEOHKCGNNSAPREAGIOPSLCLVGRGTSADHCVQ 65  
Db 12 GALGLCLAAPR-----KNVRWCTISQPEWFKCRWRMKLG-AP SITCVRRFALECI R 66  
QY 66 LIAAQEADAITLDGGAIEYACKE-HGLKPVVGEVY--DOEVGTSYVAVVVRSSHVITD 122  
Db 67 AIAEKKADAVTLGDMVFEGARDPYKURPVAEITYGTKESTQTHYVAVVVKSGNSFQID 126  
QY 123 TLKGYSCHTGINRTVGNVNPVGLVSVESGRLSVMGCDVLK----AVSDYFGSGVPGAGE 178  
Db 127 QLQGRKSCHTGLGRSAGWIIPMGIL----RPLVSWTESLEPLQGA VAKFFSASCVCPCIDR 182  
QY 179 TSYSESICRLCRGDSGSGEVCDKSPLERYDYSGAFRCIAGAGDAVFKVHSTVLENTDG 238  
Db 183 QAY-PNLQCLCKGE--GENQCACSRBEPYFGYGAFCIQLQDAGDAVAFVKETTVFEN--- 236  
QY 239 KTLPSWGQALLSQDFELLCRDGSRADVTVEQCHLARVPAHAVVVR--DTDGGGLIFRLN 297  
Db 237 --LPEKAD---RDQYELLCNLNNSRAPVDAFKCHLAQVPSHAVVARSVDGKEDLIWKLS 291  
QY 298 EQGRLF-SHEGSSPFMFSSSEAYGQKDLIFKDSITSELVPIATQTYEA-WLGEHYLHAMKGL 355  
Db 292 KAQEFKGNKRSQI LFGSP-P-GQRDILLFKDSALGLFLRIPSKVDSALYVLSRYLLTKNL 350  
QY 356 LCDNRLP-P-VLR--WCVLSTPTEIQCGDMAVAFRRQRLKPEIQCVSAKSPQHMERIOA 412  
Db 351 RETAEVKARYTRVVMCAVGPGEQKCCQWS-----QOSQONVTCATASTTDDCIVLVLK 405  
QY 413 EQVDVLTSGEDIYTAGKKYGLVPAAGEHYAPEDSSN-----SYVVAVVVRDSSH 463

Db 406 GEADALNDGGYIYTAG-KGGLVPLAENRKSXSHSLDCLVLRPTGGLYAVVVK-ANE 463  
 QY 464 AFTLDELGRKRSCHAGFGSPAGWVPVGLI-ORGPIPKDCDVLTAVGEFFNASCVPVN 522  
 Db 464 GLTWSLKKKSKCHTAVDRTAGNIPMGLIVNTG-----SC-----AFDEFFSQSCAPGA 514  
 QY 523 NPKNYSSICALCVGEQGRNKCVCNQSQRYYGYRGAFCRLVFNAGDVAFVRHTTVFDNT 582  
 Db 515 DPR--SRLCALCAGDDGLDKCVPNSKEKYGYTGAFRCLEADVDGVAFKNDTWENT 571  
 QY 583 NGHNSPWAELRSEDEYLLCPNGARAEVYSQFAACNLAQIPPHAVMVRPDNTIFTYVGLL 642  
 Db 572 NGESTADWAKNLRNDRFLCLDGTGRKPVTEAQSCHLAVAPNHAVVRSRDRAHVQVLL 631  
 QY 643 DKAQDLFGDDHKNKG-----FMFDSNNYHGQDLLFKDATVRAVPVGEKTTVYRGWGLGD 696  
 Db 632 HQ-QALFG---KNGKNCPCFKCLFKSET---KNLLFNNDTECLAKLGGRPYEEVLTGE 683  
 QY 697 YVAALGEMSSQCGSAAAPAGAPLL 722  
 Db 684 YVTAIANL--KKCS-----TSPLL 700

RESULT 8  
 TFRBP

transferrin precursor - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text change 22-Jun-1999  
 C:Accession: S16246; A61239; C61573; S00335; S02694; A26504; S14853  
 R:Banfield, D.K.; Chow, B.K.C.; Funk, W.D.; Robertson, K.A.; Umelas, T.M.; Woodworth, R.  
 Biochem. Biophys. Acta 1089, 262-265, 1991  
 A:Title: The nucleotide sequence of rabbit liver transferrin cDNA.  
 A:Reference number: S16246; MUID:91274362; PMID:2054387  
 A:Accession: S16246  
 A:Molecule type: mRNA  
 A:Residues: 1-694 <BAN>  
 A:Cross-references: EMBL:X58533; NID:g1750; PIDN:CAA41424.1; PID:g1751  
 R:Pierpaoli, W.; Dall'Ara, A.; Ya, C.; Neri, P.; Santucci, A.; Choay, J.  
 Cell. Immunol. 134, 225-234, 1991  
 A:Title: Iron carrier proteins facilitate engraftment of allogeneic bone marrow and endothelial cells.  
 A:Reference number: A61239; MUID:91191584; PMID:2013104  
 A:Accession: A61239  
 A:Molecule type: protein  
 A:Residues: 19-36 <PIE>  
 R:Chung, M.C.M.; Chan, S.L.; Shimizu, S.  
 Int. J. Biochem. 23, 609-616, 1991  
 A:Title: Purification of transferrins and lactoferrin using DEAE Affi-Gel Blue.  
 A:Reference number: A61573; MUID:91293379; PMID:2065820  
 A:Accession: C61573  
 A:Molecule type: protein  
 A:Residues: 19-26,'X',28-36,'X',38-53 <CHU>  
 R:Godovac-Zimmermann, J.  
 Biol. Chem. Hoppe-Seyler 369, 93-96, 1988  
 A:Title: Isolation, characterization and N-terminal amino-acid sequence of rabbit transferrin.  
 A:Reference number: S00335; MUID:88209278; PMID:3365331  
 A:Accession: S00335  
 A:Molecule type: protein  
 A:Residues: 19-45,'S',47-48,'Y',50 <GOD>  
 R:Evans, R.W.; Aitken, A.; Patel, K.J.  
 FEBS Lett. 238, 39-42, 1988  
 A:Title: Evidence for a single glycan moiety in rabbit serum transferrin and location of the N-glycanase cleavage site.  
 A:Reference number: S02694; MUID:89005676; PMID:3169252  
 A:Accession: S02694  
 A:Molecule type: protein  
 A:Residues: 482-515,'V',517-544 <EVA>  
 A:Note: S16-11e was also found  
 R:Heaphy, S.; Williams, J.  
 Biochem. J. 205, 611-617, 1982  
 A:Title: The preparation and partial characterization of N-terminal and C-terminal iron-transferrin - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 11-May-2000  
 C:Accession: S01384; A60520; A61573

A:Residues: 19-24,'N',26,'X',28-29,'S' <HEA>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein; iron transport; metal binding; plasma  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-694/Product: transferrin #status experimental <MAT>  
 F:19-349/Domain: transferrin repeat homology <TRH1>  
 F:355-682/Domain: transferrin repeat homology <TRH2>  
 F:27-66,37-57,136-212,155-349,176-192,179-195,189-197,245-259,357-611,363-395,373-386,4  
 F:508/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 31-2%; Score 1228; DB 1; Length 694;  
 Best Local Similarity 41.1%; Pred. No. 2.1e-85;  
 Matches 303; Conservative 109; Mismatches 225; Indels 100; Gaps 34;

QY 7 ALWLLIALTLVLGGNEVRWCATSDPEQHKCGN-----MSEAFREAGIQPSLLCVRGTSADH 62  
 Db 12 ALGLCLAVTE---KTVRCVANDHEASKANFRSMKKVLPEDG--PRIICVKKASYLD 65  
 QY 63 CVQLIAAQADAITLDGGAIYEAG-KEHGLKPVWGEVYDQEVG--TSYYAVAVVRSSHV 119  
 Db 66 CIKATAAHEADAVTLTDLGLVHEAGLTNNLKPVVAEFYGSKENPKTFYVAVALVKGSNF 125  
 QY 120 TIDTLKGVKSCHTGINTVGMVVPVGVLYVESGRSLVWGC-----VLKAVSDYEGGSC 172  
 Db 126 QNLELQKKKSCHTGLRSAGMNIPIGLY-----CDLPEPRKPLEKAVASFFSGSC 176  
 QY 173 VPGAGETSYSESLCRLCRDSSGEGVCDKSPLEYYDYSCAFRCLEAGDVAFAVRHSTV 232  
 Db 177 VPCADGADFPQ-LCQLCPG-----CGCSSVQVPFYGYSGAFKCLKGLGDVAFVQKETI 228  
 QY 233 LENTDGKTLPSWQALLSODFELLCRDGSRAVDTEWRQCHLARVPAAHVVRRA-DTDDGGL 291  
 Db 229 FEN-----LPSKDE---RDQYELLCLDNTKPKVDEYEQCHLARVPAAHVVRSDVKEDL 280  
 QY 292 IFRLLNEGQRLFSHGS-SFOMFSEAYCQKLLFKDSTSELVPIATQ-TYENWLGHYEL 349  
 Db 281 IWELLNQAEHFGKDKSGDFQLFSS-PHG-KNLLFKDSAYGFKVPPRMDANLYLGYEV 338  
 QY 350 HAMKGL---LCDPNRLP---PYLRWCVLSTPEIQKCGDMVAFRRQRLKPEIQCVSAKSP 403  
 Db 339 TAVNLRREGIC-PDPLQDECKAVKWCALSHHERLKCDEWSVTSGGL-----TECESAETP 392  
 QY 404 QHMERIOAEQVDVAVTLGSDIYTAGKYGVLPAAGEHY-----APEDSSNYYVAV 456  
 Db 393 EDCIAKIMNGEADAMSLDGGYVYIAG-QCGLVPLAENYESTDCKAPEE---GYLSVAV 448  
 QY 457 VPRDSSHAFTLDELGRKRSCHAGFGSPAGWVPVGVGLIQRGPIRKCDCDVLTAVERSEFNA 516  
 Db 449 VKK-SNPNDINWNKLGSKSCHTAVDRTAGWNIPMGLL-----YNRINHC-----RFDEFFRQ 499  
 QY 517 SCVPVNNPNKVPSSILCALCVGDEQGRNKCVCNQSQRYYGYRGAFCRLVENAGDVAFVRHT 576  
 Db 500 GCAP-GSQKN--SSLCCLCIGP---SVCAPNNREGYGYTGAFCRLVEK-GDVAFAVKSQ 551  
 QY 577 TVFDNTNGHNSPWAELRSEDEYLLCPNGARAEVYSQFAACNLAQIPPHAVMVRPDNTIF 636  
 Db 552 TVLQNTGGRNSBPWAKLKEEDFELLCLDGTGRKPVSEAHNCHLAKAPNAHVAVSRKDKAAC 611  
 QY 637 TVYGLLD---KAQDLFGDDHKNKGFMFDSNNYHGQDLLFKDATVRAVPVGEKTTVYRGWL 693  
 Db 612 VKQKLLDLQVEYGVNTVADCSKK--FCMFHSTK---KDLLFRDDTKCLVDLRGKNTYKYL 666  
 QY 694 GLDYVAALGEMSSQOCS 710  
 Db 667 GADYIKAVSNL--RKCS 681

RESULT 9

S01384

transferrin - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 11-May-2000

C:Accession: S01384; A60520; A61573



|     |  |     |
|-----|--|-----|
| 231 | ENLPDKA-----DQDYELLCKONTRRPVDYENCYLAQVPSHAVVARSVDGKDL          | 281 |
| 292 | IFRLNNGQRULFSHEGSS-FQMFSSBAYQGXDLFLKJDSSTSELVPIATO-TYBAMLGHEYL | 349 |
| 282 | IWELNQAQENFGKDGKAEQFLFSS-SHG-KDLLFTDACLGLRVPVPPKMDAKLYLGYEYF   | 339 |
| 350 | HAMKGL-LCDPNRLPPVYLRCWVLTSPETQKCGDMAVAFRRQRLKPEI-QCVSAKSPQHCM  | 407 |
| 340 | AAIQHLLRRVQGTSEPPQVMMCAVQCHETKCDMSV-----LSGGILNCNSEDWTMEDCI    | 393 |
| 408 | ERIQAEQVDVLTLSGSEDIYTAGKYGLVPAAGEHYAPEDSS-----NSGYVVA          | 455 |
| 394 | AAI AKGEGADAMSLDGGFLYTAG-KCGLVPVLAENYLSQDGKRGFGSKCVNTPEVGYVVA  | 452 |
| 456 | VVRDSSHAFPLDLBLRGKRSCHAGFGSPAGHDVVPVGLI-QRGFTRPKDCDVLTAVERSEF  | 514 |
| 453 | VVKK-SOADLTWNLSLRGKKSCHIAVGTSGAGWIIIPMGFIYNTQTSCK-----LDSEFF   | 502 |
| 515 | NASCVPVNNPNKNYPSSLCALCVGDEQGR--NKCVGNSQERYGYRGAFRCLVENAGDVAF   | 572 |
| 503 | SQSCAPGSDPE--SRLCALCSGSI SQGPAHTCAPNSHEGYHGFSGALRCLVEK-GDVAF   | 558 |
| 573 | VRHTTVFDNTNGHNSFPWAAELRSEDYELLCPNGARABVSQFAACNLAQI PHHAWMVRPD  | 632 |
| 559 | VKHPVLQNTDGRNPEAWAKDLQKEDFQLLCPDGTGRKPEVTEAQSCHLAAPVSHAVVSRKD  | 618 |
| 633 | TNI FTVGLLDKADQLFGDDHKNKGFK--MFDSSNYHGDQLLLFKDATVRAVPVGEKTTVR  | 690 |
| 619 | KADF-VRMLFNQOELEFG----RNGFEYWMFQLKFSSTEDLLFSDDTCLANLQDKIITYQ   | 673 |
| 691 | GWGLDYVAALLEGMSQOC   | 709 |
| 674 | KYLGPEYLQAIANY--RQC  | 690 |

RESULT 12  
TECHNICAL

RESULT 12  
 TFCHE  
 ovotransferrin precursor - chicken  
 N:Alternate names: conalbumin; transferrin  
 C:Species: Gallus gallus (chicken)  
 C:Date: 24-Apr-1994 #sequence revision 30-Sep-1993 #text change 22-Jun-1999  
 C:Accession: A26845; A91115; A32229; A91116; A40674; B61573; A90282; S02476; A03262  
 R:Jeltsch, J.M.; Hen, R.; Maroteaux, L.; Garnier, J.M.; Chambon, P.  
 Nucleic Acids Res. 15, 7643-7645, 1987  
 A:Title: Sequence of the chicken ovotransferrin gene.  
 A:Reference number: A26845; MUID:88015626; PMID:3658709  
 A:Accession: A26845  
 A:Molecule type: DNA  
 A:Residues: 1-705 <JEL>  
 A:Cross-references: GB:X00407; NID:G63131; PIDN:CAA68468.1; PID:G295721  
 R:Jeltsch, J.M.; Chambon, P.  
 Eur. J. Biochem. 122, 291-295, 1982  
 A:Title: The complete nucleotide sequence of the chicken ovotransferrin mRNA.  
 A:Reference number: A91115; MUID:82138851; PMID:7060577  
 A:Accession: A91115  
 A:Molecule type: mRNA  
 A:Residues: 1-82; V',84-99,'I',101-153,'W',155-238,'LN',241-685,'N',687-705 <JE2>  
 A:Cross-references: EMBL:X02009  
 A:Note: the codons given for residues 132 (AAC) and 317 (UUC) are inconsistent with the  
 R:Thibodeau, S.N.; Lee, D.C.; Palmiter, R.D.  
 J. Biol. Chem. 253, 3771-3774, 1978  
 A:Title: Identical precursors for serum transferrin and egg white conalbumin.  
 A:Reference number: A92229; MUID:78171533; PMID:649604  
 A:Accession: A92229  
 A:Molecule type: protein  
 A:Residues: 1-23 <TH>  
 R:Williams, J.; Eilenman, T.C.; Kingston, I.B.; Wilkins, A.G.; Kuhn, K.A.  
 Eur. J. Biochem. 122, 297-303, 1982  
 A:Title: The primary structure of hen ovotransferrin.  
 A:Reference number: A91116; MUID:82138852; PMID:6895872  
 A:Accession: A91116  
 A:Molecule type: protein  
 A:Residues: 1-50;54-82,'V',84-92;102-146;162-168;170-228;241-283;289-333;338-344;350-353





QY 343 ----WLGHEYLHAMKLLCDPNRLPPYLRWCVLSTPEIQKCGDMAVAFRRQRLKPEIQCV 398  
Db 332 VPRPOLCH--CHSKSAGSCPDIDAIDSAPVKWICALSHQERAKCDEWVGTNGO-----ICE 384  
QY 399 SAKSPQHCHMERIQAEQVDAVTLGSDIYTAGKKYGLVPAAGSHYAPEDSSN-----S 450  
Db 385 SAESTEDCIDKLVINGEADAMSLDGGHAYTAG--QGLVPMENYDISSCTNPQSDVFPKG 443  
QY 451 YVAVVVRDSSHAFTLDELROKRSCHAGFGSPAGWDVPVGVGALIQRGFIRPKDCDVLTA 510  
Db 444 YVAVVVRASDS--SINMNNLKKKSGCHTGVDRTAGWNIPWGLL-----FSRINHC---KF 494  
QY 511 SEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKCVCNQSERYYYGARGAFRCCLVENAGDV 570  
Db 495 DEFFSQGCAP--GYKKR--STLCDLCIGPA-----KCAPNNREGYGTGAFCQLVFK--GDV 546  
QY 571 AFVRHTTVPDNTNGHNSPEWAEALRSEDYELLCNPGARAEVSQFAACNLAQIPPHAVMVR 630  
Db 547 AFVKHQTULENTNGKNTAAWKDLQEDFOLLCPDGTGKPVTEPATCHLAQAPNHHVVR 606  
QY 631 PD--TNIFTVYGLLDPKADLF--GDDHNKNGFKMFDSSNYHGQDLLFKDQATVRAVPVGEK 686  
Db 607 KEKAARVSTV--LTAQKDLFWKDKDCTGNFCFLRSST---KDLFRDDTKLTKLPEG 660  
QY 687 TTYRGWGLDYVAALGHSQQCS 710  
Db 661 TTYEYLGAAYLQAVGNI--RKCS 682

## RESULT 15

S12100  
Transferrin precursor - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
R:Accession: S12100  
R:Moekaitis, J.E.; Pastori, R.L.; Schoenberg, D.R.  
Nucleic Acids Res. 18, 6135, 1990  
A:Title: The nucleotide sequence of Xenopus laevis transferrin mRNA.  
A:Reference number: S12100; MUID:91045087; PMID:2235503  
A:Accession: S12100  
A:Molecule type: mRNA  
A:Residues: 1-717 <MOS>  
A:Cross-references: EMBL:X54530; NID:G65158; PIDN:CAA38396.1; PID:G65159  
C:Superfamily: transferrin; transferrin repeat homology  
F:346-689/Domain: transferrin repeat homology <TRH2>

Query Match 28.7%; Score 1128.5; DB 2; Length 717;  
Best Local Similarity 37.7%; Pred. No. 8.3e-78;  
Matches 277; Conservative 106; Mismatches 252; Indels 99; Gaps 26;  
QY 22 EVRWCATSDPEQHKCKGNSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEAADAILDGA 81  
Db 25 QVRWCVKNSSELKCKDLVDTCKNKEIK--LSCVEKSNTECSLLFRKTMQMFVMTGDD 82  
QY 82 IYEAG--KEHGLKPVVGEVDQ--EVGTSYAVAVVRRSSHVTIDTLKGVSCHTGINRTV 138  
Db 83 VYKSLQPNLKPIMAEYNGSHSTETDTCYAVAVVVKSGSKFTFDELKOKKSCHTGIGKTA 142  
QY 139 GNNVPVGYLVESGRSLVSMGCDV---LKAVDYFGSCVPAGETSYSBLCRLCRGDSGG 195  
Db 143 GNNIIIGULLERKLLKWAGPDSSETWRNAVSKFKAFCVPGAKEPKLSQ-----LCAGIK-- 196  
QY 196 EGVCDKSPLEYYDYSAGFRCIAEGAGDVAFVKHSTVLTDGKTLPSWQALLSQDFEL 255  
Db 197 EHKCSRSNNEPYNYAVAGAFKCLQDDQGDVAFVKQSTVPEE-----PHKDYEL 243  
QY 256 LCDGSRADVTWRCHLARVPAHAVVVRADTDGGL--IFRLINEGQ-----RLF-----SH 305  
Db 244 LCPDNRKSIKEYKCNLAUKPAHAHLTRGRDDKSKDIIIEFLQEAQKQCECKLRLPGWG 303  
QY 306 EGSSQFMFSSEA-----YQKDLFLFKDSTSELVPIATQTYEAWLGHCHAMK--GLLCD 358

Db 304 KGSNFQQRSESYSPPIFYQ-----FSVPRSL-----FQCIQALKEGVKED 346  
QY 359 PNRLPPYLRWCVLSTPEIQKCGDMAVAFRRQRLKPEIQCVSAKSPQHCHMERIQAEQVDAV 418  
Db 347 DSAAQVKVRWCTQSKAERTKCDWTTI-----SGGAIECTEASTAEECIVQILKGDADAV 401  
QY 419 TLSGEDIYTAGKKYGLVPAAGSHYAPED-----SSNSYVAVVVRDSSHAFTLD 468  
Db 402 TLDGGMVYTAG--LCGLVPMGEYDQDDLTPCQSCSCQAKGVYVAVAVIKKGTQVWS-- 458  
QY 469 ELRGKRSCHAGFGSPAGWDVPVGVGALIQRGFIRPKDCDVLTA VSEFFNASCVPVNNPKNYP 528  
Db 459 NLRGVKTCHTAVGRTAGWNIPVGLITS-----ETANCDFAVYGE-----SCAPGSDVK--- 507  
QY 529 SSLCALCVGD-----EQGRNKCVCNQSERYYYGARGAFRCCLVENAGDVAFVRHTTVPDNTNG 584  
Db 508 SNLCALCIGDPPEKLSEREKCKSPSASEAYYGSAGFRCLVEK--GOVGFAKHTTVPENTDG 566  
QY 585 HNSEPMAAELRSEDYELLCNPGARAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDK 644  
Db 567 KNPAGWAKDLKSEDFELLCPDGSRAPVTDYKRCNLAEVPAHAHVTLTPDKR--EQVAKI VVN 625  
QY 645 AQDLFG--DDHNKNGFKMFDSSNYHGQDLLFKDQATVRAVPVGEKTTYRGWGLDYVAALG 703  
Db 626 QOSLYGRKGFQKDFQMFQSTG--GKDLLFKDSTOCLLEIPSKTTMQEFLGDKYHTAVTS 683  
QY 704 MSSQOCSGAAAPAP 717  
Db 684 LNKCSTSNEASWLP 697

Search completed: May 14, 2004, 09:43:58  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 09:36:02 ; Search time 17 Seconds  
(without alignments)  
2260.457 Million cell updates/sec

Title: US-10-049-957-4

Perfect score: 3936

Sequence: 1 MRGPSALWLLALRTVLGG.....APLLPLLPALAAALLPPAL 738

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | ID           | Description         |
|------------|--------|-------------|--------|--------------|---------------------|
| 1          | 3936   | 100.0       | 738    | 1 TRFM HUMAN | P08582 homo sapien  |
| 2          | 3300   | 83.8        | 738    | 1 TRFM MOUSE | Q9r0r1 mus musculus |
| 3          | 1308   | 33.2        | 690    | 1 TRFE ORYLA | P79819 oryzias lat  |
| 4          | 1274   | 32.4        | 691    | 1 TRFE SALSA | P80429 salmo salar  |
| 5          | 1272.5 | 32.3        | 690    | 1 TRF1 SALSA | P80426 salmo salar  |
| 6          | 1269.5 | 32.3        | 698    | 1 TRFE HUMAN | P02787 homo sapien  |
| 7          | 1263.5 | 32.1        | 706    | 1 TRFE HORSE | P27425 equus caball |
| 8          | 1261   | 32.0        | 695    | 1 TRFL HORSE | O77811 equus caball |
| 9          | 1260.5 | 32.0        | 711    | 1 TRFL HUMAN | P02788 homo sapien  |
| 10         | 1258   | 32.0        | 708    | 1 TRFL CAPHI | Q29477 capra hircu  |
| 11         | 1249   | 31.7        | 708    | 1 TRFL BOVIN | P24627 bos taurus   |
| 12         | 1249   | 31.7        | 708    | 1 TRFL CAMDR | Q9tum0 camelus dro  |
| 13         | 1248   | 31.7        | 687    | 1 TRFE ONCKI | P79815 oncorhynch   |
| 14         | 1240   | 31.5        | 695    | 1 TRFE RABIT | P19134 oryctolagus  |
| 15         | 1236.5 | 31.4        | 695    | 1 TRFE PAROL | O93429 parabulicthy |
| 16         | 1229   | 31.2        | 708    | 1 TRFL BUBBU | O77698 bubalus bub  |
| 17         | 1225.5 | 31.1        | 704    | 1 TRFE BOVIN | Q29443 bos taurus   |
| 18         | 1224.5 | 31.1        | 704    | 1 TRFL PIG   | P14632 sus scrofa   |
| 19         | 1213.5 | 30.8        | 696    | 1 TRFE PIG   | P09371 sus scrofa   |
| 20         | 1205.5 | 30.6        | 701    | 1 TRFE XENLA | Q20333 xenopus lae  |
| 21         | 1202.5 | 30.6        | 704    | 1 ICA PIG    | Q29545 sus scrofa   |
| 22         | 1202.5 | 30.6        | 705    | 1 TRFE CHICK | P02789 gallus gall  |
| 23         | 1199.5 | 30.5        | 707    | 1 TRFL MOUSE | P08071 mus musculus |
| 24         | 1196   | 30.4        | 697    | 1 TRFE MOUSE | Q92111 mus musculus |
| 25         | 1190.5 | 30.2        | 698    | 1 TRFE RAT   | P12346 rattus norv  |
| 26         | 1163   | 29.5        | 642    | 1 TRFE GADMO | Q92079 gadus morhu  |
| 27         | 1156   | 29.4        | 686    | 1 TRFE ANAPL | P56410 anas platyr  |
| 28         | 964.5  | 24.5        | 844    | 1 SAX RANCA  | P31326 rana catesb  |
| 29         | 824    | 20.9        | 726    | 1 TRF_BLADI  | Q02942 blaberus di  |
| 30         | 551.5  | 14.0        | 681    | 1 TRF_MANSE  | P22297 manduca sex  |
| 31         | 365    | 9.3         | 629    | 1 TRF_SARPE  | Q26643 sarcophaga   |
| 32         | 126    | 3.2         | 1639   | 1 LMGI DROME | P15215 drosophila   |
| 33         | 111.5  | 2.8         | 640    | 1 RAEP_CANAL | O93831 candida alb  |

#### RESULT 1

| ID | TRFM_HUMAN   | STANDARD; | PRT; | 738 AA. |
|----|--|-----------|------|---------|
| DT | P08582; Q9BOE2;  |           |      |         |
| DT | 01-AUG-1988 (Rel. 08, Created)   |           |      |         |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update)  |           |      |         |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update)  |           |      |         |
| DE | Melanotransferrin precursor (Melanoma-associated antigen p97) (CD228 antigen)  |           |      |         |
| DE | antigen)   |           |      |         |
| GN | MF12 OR MAP97.   |           |      |         |
| OS | Homo sapiens (Human)   |           |      |         |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |           |      |         |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.   |           |      |         |
| OX | NCBI_TaxID=9606;   |           |      |         |
| RN | [1]  |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM 1).  |           |      |         |
| RC | TISSUE=Melanoma;   |           |      |         |
| RX | MEDLINE=86149285; PubMed=2419904;  |           |      |         |
| RA | Rose T.M., Plozman G.D., Teplow D.B., Dreyer W.J., Hellstroem K.E., Brown J.P.;  |           |      |         |
| RT | "Primary structure of the human melanoma-associated antigen p97 (melanotransferrin) deduced from the mRNA sequence.";  |           |      |         |
| RL | Proc. Natl. Acad. Sci. U.S.A. 83:1261-1265(1986).  |           |      |         |
| RN | [2]  |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM 2).  |           |      |         |
| RC | TISSUE=Skin, and Uterus;   |           |      |         |
| RX | MEDLINE=22388257; PubMed=12477932;   |           |      |         |
| RA | Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G., Shenmen C.M., Schuler G.D., Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Moore T., Max S.I., Wang J., Hsieh F., Hopkins R.F., Jordan H., Moore T., Farmer A.A., Rubin G.M., Hong L., Diatchenko L., Marusina K., Farmer A.A., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; |           |      |         |
| RT | "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";   |           |      |         |
| RL | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  |           |      |         |
| RN | [3]  |           |      |         |
| RP | GPI-ANCHOR.  |           |      |         |
| RX | MEDLINE=94132080; PubMed=8300636;  |           |      |         |
| RA | Food M.R., Rothenberger S., Gabathuler R., Haidl I.D., Reid G., Jefferies W.A.;  |           |      |         |
| RT | "Transport and expression in human melanomas of a transferrin-like glycosylphosphatidylinositol-anchored protein.";  |           |      |         |
| RL | J. Biol. Chem. 269:3034-3040(1994).  |           |      |         |
| RN | [4]  |           |      |         |

|    |       |     |      |   |            |
|----|-------|-----|------|---|------------|
| 34 | 111   | 2.8 | 474  | 1 | GLGA_XANAC |
| 35 | 110.5 | 2.8 | 419  | 1 | PROA_BORPE |
| 36 | 107   | 2.7 | 515  | 1 | PDI_WHEAT  |
| 37 | 107   | 2.7 | 899  | 1 | ZCH2_HUMAN |
| 38 | 104.5 | 2.7 | 513  | 1 | PDI_HORVU  |
| 39 | 104   | 2.6 | 756  | 1 | PIDI_MOUSE |
| 40 | 102   | 2.6 | 1455 | 1 | FACA_HUMAN |
| 41 | 99    | 2.5 | 929  | 1 | SYA_HALN1  |
| 42 | 99    | 2.5 | 3487 | 1 | CSM2_HUMAN |
| 43 | 98.5  | 2.5 | 655  | 1 | HGFA_HUMAN |
| 44 | 97.5  | 2.5 | 650  | 1 | GYBB_MYCPN |
| 45 | 97.5  | 2.5 | 2511 | 1 | FAS_CHECK  |

#### ALIGNMENTS

RP FUNCTION.  
RX MEDLINE=96016189; PubMed=7556058;  
RA Kennard M.L., Richardson D.R., Gabathuler R., Ponka P.,  
RA Jefferies W.A.;  
RT "A novel iron uptake mechanism mediated by GPI-anchored human p97."; [5]  
RL EMBO J. 14:4178-4186 (1995).  
RN  
RP IRON-BINDING.  
RX MEDLINE=92183868; PubMed=1544447;  
RA Baker E.N., Baker H.M., Smith C.A., Stebbins M.R., Kahn M.,  
RA Hellstroem K.E., Hellstroem I.;  
RT "Human melanotransferrin (p97) has only one functional iron-binding site."; [6]  
RL FEBS Lett. 298:215-218 (1992).  
RN  
RP 3D-STRUCTURE MODELING.  
RX MEDLINE=92339524; PubMed=1633859;  
RA Garrat R.C., Jhoti H.;  
RT "A molecular model for the tumour-associated antigen, p97, suggests a Zn-binding function."; [7]  
RL FEBS Lett. 305:55-61 (1992).  
CC  
CC -!- FUNCTION: Involved in iron cellular uptake. Seems to be internalized and then recycled back to the cell membrane. Binds a single atom of iron per subunit. Could also bind zinc.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=P08582-1; Sequences=Displayed;  
CC Name=2;  
CC IsoId=P08582-2; Sequences=VSP 006557, VSP 006558;  
CC Notes-No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Found predominantly in human melanomas and in certain fetal tissues; also found in liver, epithelium, umbilical chord, placenta and sweat gland ducts.  
CC -!- DOMAIN: Composed of two homologous domains.  
CC -!- SIMILARITY: Belongs to the transferrin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; M12154; AA55992.1; -;  
DR EMBL; A00127; CAA00012.1; -;  
DR EMBL; BC001875; AAH01875.1; -;  
DR EMBL; BC002623; AAH02623.1; -;  
DR EMBL; BC007550; AAH07550.1; -;  
DR PIR; A23814; TFHUM.  
DR HSP; P19134; 1TFD.  
DR Genew; HGNC:7037; MF12.  
DR MIM; 155750; -;  
DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
DR GO; GO:0005506; F:iron ion binding; TAS.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF04005; transferrin; 2.  
DR PRINTS; PR00422; TRANSFERRIN.  
DR SMART; SM00094; TR FER; 2.  
DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
KW Transport; Iron transport; Glycoprotein; Lipoprotein; Metal-binding; Repeat; Signal; GPI-anchor; Membrane; Zinc; Alternative splicing.  
FT SIGNAL 1 19  
FT CHAIN 20 709 MELANOTRANSFERRIN (POTENTIAL).  
FT PROPEP 710 738 REMOVED IN MATURE FORM (POTENTIAL).  
FT REPEAT 20 361 1.  
FT REPEAT 362 713 2.  
FT DISULFID 26 63 BY SIMILARITY.  
FT DISULFID 36 54 BY SIMILARITY.

|                           |   |     |   |
|---------------------------|---|-----|---|
| FT DISULFID               | 130   | 216 | BY SIMILARITY.  |
| FT DISULFID               | 172   | 189 | BY SIMILARITY.  |
| FT DISULFID               | 186   | 199 | BY SIMILARITY.  |
| FT DISULFID               | 257   | 271 | BY SIMILARITY.  |
| FT METAL                  | 78  | 78  | IRON 1 (BY SIMILARITY).   |
| FT METAL                  | 107   | 107 | IRON 1 (BY SIMILARITY).   |
| FT METAL                  | 210   | 210 | IRON 1 (BY SIMILARITY).   |
| FT METAL                  | 279   | 279 | IRON 1 (BY SIMILARITY).   |
| FT METAL                  | 421   | 421 | IRON 2 (BY SIMILARITY).   |
| FT METAL                  | 451   | 451 | IRON 2 (BY SIMILARITY).   |
| FT METAL                  | 556   | 556 | IRON 2 (BY SIMILARITY).   |
| FT METAL                  | 625   | 625 | IRON 2 (BY SIMILARITY).   |
| FT BINDING                | 132   | 132 | CARBONATE 1 (BY SIMILARITY).  |
| FT BINDING                | 136   | 136 | CARBONATE 1 (BY SIMILARITY).  |
| FT BINDING                | 138   | 138 | CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY).   |
| FT BINDING                | 139   | 139 | CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY).   |
| FT LIPID                  | 709   | 709 | GPI-anchor amidated cysteine (Potential).   |
| FT CARBOHYD               | 38  | 38  | N-LINKED (GLCNAC. .) (POTENTIAL).   |
| FT CARBOHYD               | 135   | 135 | N-LINKED (GLCNAC. .) (POTENTIAL).   |
| FT CARBOHYD               | 515   | 515 | N-LINKED (GLCNAC. .) (POTENTIAL).   |
| FT VARSPLIC               | 238   | 302 | KTLPSWGQALLSQDFELLCRDGSRADVTWQCHLARVPA<br>HAVVVRADTGGGLIFRLNEGQL -> ESPSRQTWRSE<br>EEEGCPAHEARTRMRSSAGQAKWAPVHRPQDESCKGEF<br>GKRAKSRDMLG (in isoform 2).<br>/FTId=VSP 006557.<br>Missing (in isoform 2).<br>/FTId=VSP 006558. |
| FT VARSPLIC               | 303   | 738 |   |
| FT SEQUENCE               | 738 AA; 80241 MW; 6E6086E894D7B955 CRC64;                   |     |   |
| Query Match               | 100.0%; Score 3936; DB 1; Length 738;                       |     |   |
| Best Local Similarity     | 100.0%; Pred. No. 1.7e-292;                                 |     |   |
| Matches 738; Conservative | 0; Mismatches 0; Indels 0; Gaps 0;                          |     |   |
| QY 1                      | MRGPGALWLLALRTVLGGMVEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTS  | 60  |   |
| DB 1                      | MRGPGALWLLALRTVLGGMVEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTS  | 60  |   |
| QY 61                     | DHCVOLIAAQADAITLGGAIYEAGKEHGLKPVVGEYDQEVGTSYVAVVRRSSHT      | 120 |   |
| DB 61                     | DHCVOLIAAQADAITLGGAIYEAGKEHGLKPVVGEYDQEVGTSYVAVVRRSSHT      | 120 |   |
| QY 121                    | IDTLKGVSCHTGINTRTGWNVPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPGAGETS | 180 |   |
| DB 121                    | IDTLKGVSCHTGINTRTGWNVPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPGAGETS | 180 |   |
| QY 181                    | YSESLCRLCRGDSGEGVCDKSPLEYYDYSGAFRLAEGADVAFVKHSTVLENTDGT     | 240 |   |
| DB 181                    | YSESLCRLCRGDSGEGVCDKSPLEYYDYSGAFRLAEGADVAFVKHSTVLENTDGT     | 240 |   |
| QY 241                    | LPSWGQALLSQDFELLCRDGSRADVTWQCHLARVPAHAVVVRADTGGGLIFRLNEGQ   | 300 |   |
| DB 241                    | LPSWGQALLSQDFELLCRDGSRADVTWQCHLARVPAHAVVVRADTGGGLIFRLNEGQ   | 300 |   |
| QY 301                    | RLFSHEGSSGFOMFSSEAYGQDLPFKDSTSELVPIATQTYEAWLGHEYLHAKGLLCDPN | 360 |   |
| DB 301                    | RLFSHEGSSGFOMFSSEAYGQDLPFKDSTSELVPIATQTYEAWLGHEYLHAKGLLCDPN | 360 |   |
| QY 361                    | RLPYLRWCVLSTPEIQKCGDMAVAFRRQRKPEIQCVSAKSPOCHMERIOAEQVDVTL   | 420 |   |
| DB 361                    | RLPYLRWCVLSTPEIQKCGDMAVAFRRQRKPEIQCVSAKSPOCHMERIOAEQVDVTL   | 420 |   |
| QY 421                    | SGEDIYTAGKYGLVPAAGEHYAPEDSSNSYVAVVVRDRSSHATLDELGRKRSCHAGF   | 480 |   |
| DB 421                    | SGEDIYTAGKYGLVPAAGEHYAPEDSSNSYVAVVVRDRSSHATLDELGRKRSCHAGF   | 480 |   |
| QY 481                    | GSPAGWDVPVGCALIORGIRPKDCDLTAVSEFFNASCVPVNNPKVNPSSLCALCVGBEQ | 540 |   |
| DB 481                    | GSPAGWDVPVGCALIORGIRPKDCDLTAVSEFFNASCVPVNNPKVNPSSLCALCVGBEQ | 540 |   |
| QY 541                    | GRNCKVGNQBRYGYRGCAFRLVENAGDVAFVRHTTVFDNTNGHNSPFWAAELSEDEYE  | 600 |   |
| DB 541                    | GRNCKVGNQBRYGYRGCAFRLVENAGDVAFVRHTTVFDNTNGHNSPFWAAELSEDEYE  | 600 |   |



Db 1 MRLSVTFLLSLRTVVCVMEVQWCTISDAEQKCKMDSEAFQAGIRPSLLCVQNSA 60  
 Qy 61 DHCYQLIAQAEADAITLDGAIYEAGKEHGLKPVVGEVYDQVCTSYAYAVVRRSSHVT 120  
 Db 61 DHCYQLIKEQADAITLDGAIYEAGKEHGLKPVVGEVYDQVCTSYAYAVVRRSNVT 120  
 Qy 121 IDTLKGVKCHTGINRTVGNVPGVYLVESGRSLVWGCVDLVKAVSDYFGGSCVPGAGETS 180  
 Db 121 INTLKGKVSCHTGINRTVGNVPGVYLVESGHLVWGCVDLVKAVSDYFGGSCVPGAGETS 180  
 Qy 181 YSESLCRLCRGSSGEGVCDKSPLEKYDYDYGAFRLAEGAGDVAFVXHSVLENTDGT 240  
 Db 181 HSESLCRLCRGSSGHNVCDSKPLEKYDYDYGAFRLAEGAGDVAFVXHSVLENTDGT 240  
 Qy 241 LPSWGALLSQPELLCRGSRADVTWFOCHLARPAPAHAVVVRADTGGGLIFRLNEQ 300  
 Db 241 LPSWGKLSMSEDFOLLCRGSRADITWRRCHLAKVPAHAVVVRGMDGGLIFOLLNEQ 300  
 Qy 301 RLFSHEGSSFOFSSKAYSGKDLFPKSTSELVPIATQTYEAWLGHVYHAKMGLLCPN 360  
 Db 301 RLFSHEDSSFOFSSKAYSGKDLFPKSTSELVPIATQTYEAWLGOEYLAQWGLLCPN 360  
 Qy 361 RLPLYLWCVLSTPEIKQKGMMAVAFRRQRLKPEIQCVSAKSPQCHMERIQAEQVDAVTL 420  
 Db 361 RLPLYLWCVLSAPEIQKGMMAVAFRRQRLKPEIQCVSAESPEHCMEIQAGHTDAVTL 420  
 Qy 421 SEDIYTAGKTKGLVPAAGEHAPEDSSNSYVAVVVRDSSHAFTLDELGRKRSCHAGF 480  
 Db 421 RGEYIRAGKTKGLVPAAGEHAPEDSSNSYVAVVVRDSSHAFTLDELGRKRSCHPYL 480  
 Qy 481 GSPAGWDVPVGLIQRGFRPKDCDVLTAVSFFNASCVPVNNKPNYSSLCALCVGDEO 540  
 Db 481 GSPAGWEVPIGSLIQRGFRPKDCDVLTAVSFFNASCVPVNNKPNYSSLCALCVGDEK 540  
 Qy 541 GRNKCVCNSQERYGGRGAFRLCVENAGDVAFVRRHTTVDNTNGHNSPAAELRSEDYE 600  
 Db 541 GRNKCVCNSQERYGGRGAFRLCVENAGDVAFVRRHTTVDNTNGHNSPAAELRSEDYE 600  
 Qy 601 LLCPNGAREVSOFAACNLAIQIPPHAVVVRPDTNIFTVYGLLDKADLFGDDHKNKGFQ 660  
 Db 601 LLCPNGAREVSOFAACNLAIQIPPHAVVVRPDTNIFTVYGLLDKADLFGDDHKNKGFQ 660  
 Qy 661 FDSSNVHGDLFPKATRAVPVGEKTYRGMGLDYVAALRGMSQCSGAAAPAGAP 720  
 Db 661 FDSSNVHGDLFPKATRAVPVGEKTYRGMGLDYVAALRGMSQCSGAAAPAGAP 720  
 Qy 721 LLPPLPALAARLLPPAL 738  
 Db 721 LLPPLPALAARLLPPAL 738

RESULT 3

ID TRFE ORYLA STANDARD; PRT; 690 AA.  
 AC P79619;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serotransferrin precursor.  
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;  
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.  
 OX NCBI\_TaxID=8090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96414927; PubMed=8817928;  
 RA Mikawa N., Hirano I., Aoki T.;  
 RT "Structure of medaka transferrin gene and its 5'-flanking region."  
 RL Mol. Mar. Biol. Biotechnol. 5:225-229 (1996).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING

CC OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: Composed of two homologous domains.  
 CC -!- SIMILARITY: Belongs to the transferrin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; D64033; BAAL0901.1; -.  
 CC HSP; P56410; IAOV.  
 CC InterPro; IPR001156; Transferrin.  
 CC Pfam; PF00405; transferrin; 2.  
 CC PRINTS; PR00422; TRANSFERRIN.  
 CC SMART; SM00094; TR\_FER; 2.  
 CC PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 CC PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 CC PROSITE; PS00207; TRANSFERRIN\_3; FALSE NEG.  
 CC KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 CC Signal.  
 CC SIGNAL 1 17  
 CC CHAIN 18 690  
 CC SEROTRANSFERRIN.  
 CC DISULFID 27 49  
 CC BY SIMILARITY.  
 CC DISULFID 126 206  
 CC BY SIMILARITY.  
 CC DISULFID 171 185  
 CC BY SIMILARITY.  
 CC DISULFID 234 248  
 CC BY SIMILARITY.  
 CC DISULFID 343 379  
 CC BY SIMILARITY.  
 CC DISULFID 353 370  
 CC BY SIMILARITY.  
 CC DISULFID 404 682  
 CC BY SIMILARITY.  
 CC DISULFID 419 643  
 CC BY SIMILARITY.  
 CC DISULFID 452 530  
 CC BY SIMILARITY.  
 CC DISULFID 476 671  
 CC BY SIMILARITY.  
 CC DISULFID 486 499  
 CC BY SIMILARITY.  
 CC DISULFID 496 513  
 CC BY SIMILARITY.  
 CC DISULFID 570 584  
 CC BY SIMILARITY.  
 CC METAL 73 73  
 CC IRON 1 (BY SIMILARITY).  
 CC METAL 103 103  
 CC IRON 1 (BY SIMILARITY).  
 CC METAL 200 200  
 CC IRON 1 (BY SIMILARITY).  
 CC METAL 256 256  
 CC IRON 1 (BY SIMILARITY).  
 CC METAL 394 394  
 CC IRON 2 (BY SIMILARITY).  
 CC METAL 429 429  
 CC IRON 2 (BY SIMILARITY).  
 CC METAL 524 524  
 CC IRON 2 (BY SIMILARITY).  
 CC METAL 592 592  
 CC CARBONATE 1 (BY SIMILARITY).  
 CC METAL 128 128  
 CC CARBONATE 1 (BY SIMILARITY).  
 CC METAL 132 132  
 CC CARBONATE 1 (BY SIMILARITY).  
 CC METAL 134 134  
 CC CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 CC SIMILARITY).  
 CC METAL 135 135  
 CC CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 CC SIMILARITY).  
 CC METAL 454 454  
 CC CARBONATE 2 (BY SIMILARITY).  
 CC METAL 458 458  
 CC CARBONATE 2 (BY SIMILARITY).  
 CC METAL 460 460  
 CC CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 CC SIMILARITY).  
 CC METAL 461 461  
 CC CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 CC SIMILARITY).  
 CC SEQUENCE 690 AA; 74600 MW; 2F996CALAE79570 CRC64;  
 Query Match 33.2%; Score 1308; DB 1; Length 690;  
 Best Local Similarity 40.3%; Pred. No. 5.8e-92;  
 Matches 300; Conservative 116; Mismatches 215; Indels 114; Gaps 29;  
 Qy 12 LALRTVLG-----GMEVWRWCATSDPEQHKCNMSEAFREAGIQPSLLCVRGTSADH 62  
 Db 4 LLLLTLLCLLAALAVPAQKWKVCQDEFRKCSDLAA-----SFAFSCVKKESTLD 57  
 Qy 63 CVQLIAQAEADAITLDGAIYEAG-KEHGLKPVVGEVYDQVCTSYAYAVVRRSSHVT 121  
 Db 58 CIIAIKAGEADAITVDGDDVYTAGLNNDLHPITAEADYGTSETCYAYAVAKKGTFTGI 117

QY 122 DTLLGVKSGCHTGINRTVMNVPVGYLVESGRLSVMGCD---VLKAVSDYDFGSCVPGAGE 178  
 Db 118 RDLRGKSGCHTGLGKSAGNIPITGLVSMDIQWAGVEDKPEVEESTFFQASCVPGATR 177  
 QY 179 TSVSESLCRLCRGDSGEGVCDKSPLERYDYDSGAFRCLEAGADVAFVKHSTVLENTDG 238  
 Db 178 GS---KLCECKGD-----CSRSKEPYDYDNGAFNCLAEAGADVAFVKHSTVLENTDG 228  
 QY 239 KTLPSWQALLSQDFELLCLRGSRADVTEWRQCHLARVPAAHVVRADTD-GGLIFRLLN 297  
 Db 229 K-----YELLCHDNTAPIDDYKKNLARVPAAHVVRADTD-GGLIFRLLN 275  
 QY 298 EGQRLFSHSGSSQMFSSAYG-QKOLLFKDSTSELVPIATQTYE-AWLGHYELHAMKGL 355  
 Db 276 SVQ-----NFNLFSEAYAPSKNLMFKDSTQRLVKLPQNTDSFLYLCAQYMSIVRS 327  
 QY 356 LCDPN--RLPPYLRCVLTSPFETQK-----GDMVAFRRQRLKPIQCVSAKSPQ 404  
 Db 328 KKSQSVGTNSNAIKWCAVGHAEAKKCDTWSINSVTDATA-----IEQNAPSVE 377  
 QY 405 HCMERIOAEQVDAVTLGSGEDYTAGKKYGLVPAAGHYAPE-----DSSNSYVVAVVR 458  
 Db 378 ECLKIMRKADAWVDGGEVYTAG-KGLVPAMVEQYDAELCSSGGQASSYAVAVVR 436  
 QY 459 RDSSHAFTLDELGRKSRCHAGFGSPAGWDVPVPGALIQRFIRPKDCDVLTAVERSEFNASC 518  
 Db 437 KDS--GVTWENLGRKSGCHTIGRTAGWNIPMG---RIYDQTKDCD---FTKFPSPGC 486  
 QY 519 VPNNPKVPSISICALCV-----GDEQRNKCVCNQSERYGYGAFCLVENAGDVA 571  
 Db 487 AP--EPK--PALHCALCKAAKLSGDEA---KCKARPEQYGYAGAFCLAEAGADVA 539  
 QY 572 FVRHTTVFNTNGHNSPWAELRSEYELLCPNGARAEVSQFAACNLAQIIPHAVMVRP 631  
 Db 540 FIKHTIVGENTDG-NQPDWARSKSDYQICPKGVPVISEYASCNLAVVPAHAVMVRP 598  
 QY 632 DTNIFTYVGLLKAQDLFGDHNKNGFKMFDSSNYHQDILLFKDAT--VRAVPVGEKTTY 689  
 Db 599 ESR-SDVVRVLQVQTFFGAGSDPSFKLFQSQN--GNLLFKDSTKCLQEVNPA--TSY 653  
 QY 690 RGMGLDYVAALGMSQSCGAA 714  
 Db 654 DQFLGSGYMEAM--TSLRKCSDTAS 676

## RESULT 4

TRF2\_SALSA STANDARD; PRT; 691 AA.  
 AC P80429;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serotransferrin II precursor (Siderophilin II) (STF II).  
 GN STF2.  
 OS Salmo salar (Atlantic salmon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.  
 OX NCBI\_TaxID=8030;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=94122797; PubMed=8293074;  
 RA Kvingedal A.M., Roervik K.A., Alestroem P.;  
 RT "Cloning and characterization of Atlantic salmon (Salmo salar) serum  
 transferrin cDNA."  
 RL Mol. Mar. Biol. Biotechnol. 2:233-238(1993).  
 RN [2]  
 RN SEQUENCE OF 19-37.  
 RC TISSUE=Serum;  
 RA Røed K.H., Dehli A.K., Fløengrud R., Midthjell L., Roervik K.A.;  
 RT "Immunoassay and partial characterization of serum transferrin from

RT Atlantic salmon (Salmo salar L.).";  
 RL Fish Shellfish Immunol. 5:71-80(1995).  
 CC -I- FUNCTION: Transferrins are iron binding transport proteins which  
 CC can bind two atoms of ferric iron in association with the binding  
 CC of an anion, usually bicarbonate. It is responsible for the  
 CC transport of iron from sites of absorption and heme degradation to  
 CC those of storage and utilization. Serum transferrin may also have  
 CC a further role in stimulating cell proliferation.  
 CC -I- SUBUNIT: Monomer.  
 CC -I- SUBCELLULAR LOCATION: Secreted.  
 CC -I- TISSUE SPECIFICITY: Abundant in liver and serum with smaller  
 CC amounts found in the stomach and kidney.  
 CC -I- DOMAIN: Composed of two homologous domains.  
 CC -I- SIMILARITY: Belongs to the transferrin family.  
 DR HSSP; P56410; LAOV.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 1.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 691 SEROTRANSFERRIN II.  
 FT DISULFID 28 50 BY SIMILARITY.  
 FT DISULFID 127 207 BY SIMILARITY.  
 FT DISULFID 172 186 BY SIMILARITY.  
 FT DISULFID 235 249 BY SIMILARITY.  
 FT DISULFID 343 379 BY SIMILARITY.  
 FT DISULFID 353 370 BY SIMILARITY.  
 FT DISULFID 404 682 BY SIMILARITY.  
 FT DISULFID 419 643 BY SIMILARITY.  
 FT DISULFID 451 530 BY SIMILARITY.  
 FT DISULFID 475 671 BY SIMILARITY.  
 FT DISULFID 485 499 BY SIMILARITY.  
 FT DISULFID 496 513 BY SIMILARITY.  
 FT DISULFID 570 584 BY SIMILARITY.  
 FT METAL 74 74 IRON 1 (BY SIMILARITY).  
 FT METAL 104 104 IRON 1 (BY SIMILARITY).  
 FT METAL 201 201 IRON 1 (BY SIMILARITY).  
 FT METAL 257 257 IRON 1 (BY SIMILARITY).  
 FT METAL 394 394 IRON 2 (BY SIMILARITY).  
 FT METAL 428 428 IRON 2 (BY SIMILARITY).  
 FT METAL 524 524 IRON 2 (BY SIMILARITY).  
 FT METAL 592 592 IRON 2 (BY SIMILARITY).  
 FT CARBOHYD 169 169 N-LINKED (GLCNAC... (POTENTIAL).  
 FT BINDING 129 129 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 134 134 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 136 136 SIMILARITY).  
 FT BINDING 137 137 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 453 453 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 457 457 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 459 459 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 460 460 SIMILARITY).  
 FT BINDING 460 460 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 691 691 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 FT SEQUENCE 691 AA; 74663 MW; 8D2431663677CF2D CRC64;  
 Query Match 32.4%; Score 1274; DB 1; Length 691;  
 Best Local Similarity 38.9%; Pred. No. 2.3e-89;  
 Matches 287; Conservative 130; Mismatches 224; Indels 96; Gaps 28;  
 QY 11 LLALRTVLGME-----VRWCATSDPEQHKCNMSEAPREAGIOPSLCVRGTS 60  
 Db 3 LLLLSALLGLCLATAYAAPAEGIVKWCQSEQLKCHDLAAKVAE-----FSCVRKDG 56  
 QY 61 DHCQVLIQAQDAITLDGGAIEAG-KEHGLKPVVGEVYDQVGVSYAVAVVRSSHV 119  
 Db 57 FECIQAIKGEADAITLDGGDIYTAGLTNYGLQPIAEDYGEDSDTCYAVAVAKGTAF 116

```
QY 120 TIDTLKGVKSCHTGINRTVGMNVPVGLVSGRLSMVGC---VLKAVSDYFGGSCVPGA 176
Db 117 GKTLRGKSGHTGKAGNIPITGLVTESQIRKAGIEDRPVESAVDSDFNASCAPGA 176
QY 177 GETSYSESLRCRGGSSGEGVCDKSPLEYYDYGAFRCIAEAGDGVAFVKHSTVLENT 236
Db 177 ---TWGSKLCOJCKGD-----CSRSHKEPYDYAGAFQCLKDGAGDGVAFIKPLAV--- 223
QY 237 DGKTLPSWGQALLSODFELLCRDGSRADVTWROCHLARVPAHAVVADTDGGLIFRL 296
Db 224 -----PRAEKASYELLCKDGTFRASIDSYKTCHLARVPAHAVVSRDPE--LANRIY 272
QY 297 NEGQRLFSHEGSSFFMFSEAVGQKDLFPKSTSELVPIATOTYE-AWLGHEYLHAMKGL 355
Db 273 NKLAV-----KDFLFSGSGVAAKMLFKSAQKLQVPTTDSFLYLGAEYMTIRSL 327
QY 356 LCD--PNRLPPYLRVCLSTPEIQCGDMV--AFRRQRLKPEIQCVSAKSPQHMERIOA 412
Db 328 KKSQATGASSRAIKCAVGAHAEKGCDDTWTINSFADGESK--ISCQDAPTVEECIKIMR 385
QY 413 EQVDVAVTSGEDIYTAGKYGLVPAAGEHY-----APEDSSNYYVAVVVRDSSHAFT 466
Db 386 KEADAIAVDGGEVYTAG-KCGLVPVVMVEQYDADLCSAPGEAS-SYAVAVAKKGS--GLT 441
QY 467 LDELGRKSGHAGFGSPAGMDVPVPCALIQRGFIRPKDCDVLTVAVSEFFENASCVP---VNN 523
Db 442 WKTLLGKRSCHTGLGTAGNIPMLGIHQ-----ETNDCD----FTKYFSKGCAPGSEVGS 493
QY 524 PKNYPSSLCALCVGDEQGR---NKCVGNSQBRYYGYRGAFRCFLVENAGDVAFVRHTTVF 579
Db 494 P-----FCAQCKGSGKARGGDEDRCKARSEQYVGYTGAFRCFLVEDAGDVAFIKHTIVP 547
QY 580 DNTNGHNSPMAELRSEDYELLCPNGARAQVSAQFAACNLAQIPPHAVVVRPDNTIFTY 639
Db 548 ESTDG-NGPDWAKLSSDFELLQDGTTPQVTKFSECHLAKVPAHAVITRPTETR-GDVV 605
QY 640 GLLDRAQDLFGDNHKNKGFMDSSNYHGODLLFKDAT--VRAVPVGEKTVYRGWLGLDY 697
Db 606 SILLELQAKFGSGSDSFRMFQSS--VERNLKFDSTKCLQEIPIKG--TKYQDFLGKEY 661
QY 698 VAALEGMSQQCSGAAA 714
Db 662 MIAMQ--SLRKCS DSTS 676

RESULT 5
ID TRF1 SALSA STANDARD; PRT; 690 AA.
AC P80426,
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serotransferrin I precursor (Siderophilin I) (STF I).
OS STF1.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94122797; PubMed=8293074;
RA Kvingedal A.M., Roervik K.A., Alestroem P.;
RT "Cloning and characterization of Atlantic salmon (Salmo salar) serum
transferrin cDNA";
RL Mol. Mar. Biol. Biotechnol. 2:233-238 (1993).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RX MEDLINE=95121925; PubMed=7821802;
RA Kvingedal A.M.;
RT "Characterization of the 5' region of the Atlantic salmon (Salmo
```

```
RT salar) transferrin-encoding gene.";
RL Gene 150:335-339 (1994).
RN [3]
RP SEQUENCE OF 19-37.
RC TISSUE=Serum;
RA Roed K.H., Dehli A.K., Flengsrud R., Midtjell L., Roervik K.A.;
RT "Immunoassay and partial characterization of serum transferrin from
Atlantic salmon (Salmo salar L.).";
RL Fish Shellfish Immunol. 5:71-80 (1995).
CC -!- FUNCTION: Transferrins are iron binding transport proteins which
can bind two atoms of ferric iron in association with the binding
of an anion, usually bicarbonate. It is responsible for the
transport of iron from sites of absorption and heme degradation to
those of storage and utilization. Serum transferrin may also have
a further role in stimulating cell proliferation.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Abundant in liver and serum with smaller
amounts found in the stomach and kidney.
CC -!- DOMAIN: Composed of two homologous domains.
CC -!- SIMILARITY: Belongs to the transferrin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L20313; AAA18838.1; -.
DR EMBL; L26909; AAC42221.1; -.
DR PIR; I51350; I51350.
DR PIR; T11749; T11749.
DR HSSP; P56410; IAOV.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 1.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
Signal.
FT SIGNAL 1 18
FT CHAIN 19 690 SEROTRANSFERRIN I.
FT DISULFID 28 50 BY SIMILARITY.
FT DISULFID 127 207 BY SIMILARITY.
FT DISULFID 172 186 BY SIMILARITY.
FT DISULFID 235 249 BY SIMILARITY.
FT DISULFID 343 379 BY SIMILARITY.
FT DISULFID 353 370 BY SIMILARITY.
FT DISULFID 404 681 BY SIMILARITY.
FT DISULFID 419 642 BY SIMILARITY.
FT DISULFID 451 529 BY SIMILARITY.
FT DISULFID 475 670 BY SIMILARITY.
FT DISULFID 485 499 BY SIMILARITY.
FT DISULFID 496 512 BY SIMILARITY.
FT DISULFID 569 583 BY SIMILARITY.
FT METAL 74 74 IRON 1 (BY SIMILARITY).
FT METAL 104 104 IRON 1 (BY SIMILARITY).
FT METAL 201 201 IRON 1 (BY SIMILARITY).
FT METAL 257 257 IRON 1 (BY SIMILARITY).
FT METAL 394 394 IRON 2 (BY SIMILARITY).
FT METAL 428 428 IRON 2 (BY SIMILARITY).
FT METAL 523 523 IRON 2 (BY SIMILARITY).
FT METAL 591 591 IRON 2 (BY SIMILARITY).
FT CARBOHYD 169 169 N-LINKED (GLCNAC...) (POTENTIAL).
FT BINDING 129 129 CARBONATE 1 (BY SIMILARITY).
FT BINDING 134 134 CARBONATE 1 (BY SIMILARITY).
FT BINDING 136 136 CARBONATE 1 (VIA AMIDE NITROGEN) (BY
SIMILARITY).
FT BINDING 137 137 CARBONATE 1 (VIA AMIDE NITROGEN) (BY
```

```
FT BINDING 453 453 SIMILARITY).
FT BINDING 457 457 CARBONATE 2 (BY SIMILARITY).
FT BINDING 459 459 CARBONATE 2 (BY SIMILARITY).
FT BINDING 460 460 CARBONATE 2 (VIA AMIDE NITROGEN) (BY
FT BINDING 460 460 CARBONATE 2 (VIA AMIDE NITROGEN) (BY
FT CONFLICT 34 34 SIMILARITY).
FT SEQUENCE 690 AA; 74597 MW; D1F4FC74A6AFA3FB CRC64;

Query Match 32.3%; Score 1272.5; DB 1; Length 690;
Best Local Similarity 39.1%; Pred. No. 3e-89;
Matches 289; Conservative 128; Mismatches 221; Indels 101; Gaps 29;

QY 11 LIALRTVLGME-----VRWCATSDPEQHKCKNMSEAFREAGIOPSLLCVGRGTS 60
DB 3 LLLLSALLGLATAYAAPAGGIVKMKVSEQLKCHDLAAKVAE-----FSCVRKDG 56
QY 61 DHCVQIIAQEADAITLDGAIYEAG-KEHGLKPVVGEVVDQEVGTSYAVAVVRSSHV 119
DB 57 FECIOAKGGEADAITLDGDIYTAGLTNYGLOPIIAEDYGEDSDTCYAVAVAKKGTAF 116
QY 120 TIDTLKGVKSGHGINRTWGVNVPVGLVSGRLSVMGCD---VLKAVSDYFGSCVPGA 176
DB 117 GFKTLRGKKSCHTGLGKSGAGNIPIGTLVTEQIRWAGIEDRPVESAIVDFNFNACPGA 176
QY 177 GETSYSESICRLCRGDSGBGVCDKSPLERYDYGAFRCIARAGDVAFAVFKHSTVLENT 236
DB 177 ---TWGSKULCQCKGD-----CSRSHKEPYDYAGAPQLCKDAGDVAFIKPLAV--- 223
QY 237 DGKTLPSWGOALLSODPELLCRDGRADVTWEMRQCHLARVPAHVVVRADTDGGLIFRLL 296
DB 224 -----PAEAKASVELLCKDGTASIDSYKTHLARVPAHVVVRKDPE--LANRIY 272
QY 297 NEGQRLFSHEGSGFQFSEAYCQKDLFLKDSITSELVPTATQTYE-AWLGHEVYHAKGL 355
DB 273 NKLAV-----KDFNLFSSDGYAAKNLMPKDSAQKLVLQPTTTDSFLYLGAETMSTIRSL 327
QY 356 LCD--PNRLPPYLRWCVLPSTPEIQKGDMAV-AFRQRRLKPEIQCVSAKSPHOCHMERIOA 412
DB 328 KKSQATGASSRAIKCAVGAHKGKCDTWTINSFADGESK--ISCQDAPTEBECIKIMR 385
QY 413 EQVDVTLGSEDITYTAGKYGLVPAAGEHY-----APEDSNSYVYVAVVRDSSHAFT 466
DB 386 KEADAIAVDGVEYTAG-KCGLVPVMVEQYDADLCSAPGEAS-SYVAVAVAKKGS--GLT 441
QY 467 LDELKRSCHAGFGSPAGWDVPVGLIQGFIRPKDCDVLTVAFSEFNASCVP-----VNN 523
DB 442 WKTLLKGRSCHTGLGRTAGNIPMGLIHQ-----ETNDCD-----FTKYFGKCAPGSEVGS 493
QY 524 PKNYPSSLCAIC-----VGDEGRNKCVCNSQERYGYRGAPRCICLVENAGDVAFVRHTT 577
DB 494 P-----FCAQCKGSKAVGDEY---RCKARSEQYGYTGAPRCILVEDAGDVAFIKHTI 544
QY 578 VFDNTNGHNSPWAELRSEDYELLCPNGARAEVSOFAACNLAQIPPHAVMVRPDTNIFT 637
DB 545 VPBSTDG-NGPDWAKDKSSDFELLQDGTQPTQVTFSECHLAKVPAHAVITRPEIR-GD 602
QY 638 VYGLLDKAQDLFDHINKGKFMFDSNTHGQDLFLFKDAT--VRAVPVGEKTYRQWLGL 695
DB 603 VVSILLELQAKFGSGSDSFRMFQSS--VEKNLLPKDSTKCIQETPKG--TKYQDFLGK 658
QY 696 DYVAALLEGMSQOCGAAA 714
DB 659 EYMIAMQ--SLRCKSDSTS 675
```

## RESULT 6

```
TRFE HUMAN
ID TRFE HUMAN STANDARD; FRG; 698 AA.
AC P02787; O43890; Q9NQ88; Q9UHV0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
```

```
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serotransferrin precursor ("Transferrin") (Siderophilin) (Beta-1-metal
GN binding globulin) (PRO1400).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS TF*B2; TF*CHI AND TF*DI.
RX MEDLINE=84194084; PubMed=6585826;
RA van Bragt P.H., Baldwin W.D., Bowman B.H.;
RA "Human transferrin: cDNA characterization and chromosomal
RT localization.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:2752-2756(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88056305; PubMed=3678832;
RA Schaeffer E., Lucero M.A., Jeltsch J.-M., Py M.-C., Levin M.J.,
RA Chambon P., Cohen G.N., Zakin M.M.;
RT "Complete structure of the human transferrin gene. Comparison with
RT analogous chicken gene and human pseudogene.";
RL Gene 56:109-116(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=923211399; PubMed=1809186;
RA Hersberger C.B., Larson J.L., Arnold B., Rosteck P.R. Jr.,
RA Williams P., Dehoff B., Dunn P., O'Neal K.L., Riemen M.W.,
RA Tice P.A.;
RT "A cloned gene for human transferrin.";
RL Ann. N.Y. Acad. Sci. 646:140-154(1991).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT ATRANSFERRINEMIA PRO-477.
RX MEDLINE=20563920; PubMed=11110675;
RA Beutler E., Gelbart T., Lee P.L., Trevino R., Fernandez M.A.,
RA Fairbanks V.F.;
RT "Molecular characterization of a case of attransferrinemia.";
RL Blood 96:4071-4074(2000).
RN [5]
RP SEQUENCE OF 99-698 FROM N.A.
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Bi J., Zhang Y., Liu M.,
RA He F.;
RT "Functional prediction of the coding sequences of 33 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RN Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RP [6]
RP SEQUENCE OF 422-698 FROM N.A.
RX MEDLINE=84153910; PubMed=6322780;
RA Uzan G., Frain M., Park I., Besmond C., Maessen G., Trepas J.S.,
RA Zakin M.M., Kahn A.;
RT "Molecular cloning and sequence analysis of cDNA for human
RT transferrin.";
RN Biochem. Biophys. Res. Commun. 119:273-281(1984).
RP [7]
RP SEQUENCE OF 20-698.
RX MEDLINE=83160878; PubMed=6833213;
RA McGilivray R.T.A., Mendez E., Shewale J.G., Sinha S.K.,
RA Lineback-Zins J., Brew K.;
RT "The primary structure of human serum transferrin. The structures of
RT seven cyanogen bromide fragments and the assembly of the complete
RT structure.";
RL J. Biol. Chem. 258:3543-3553(1983).
RN [8]
RP SEQUENCE OF 73-698 FROM N.A.
RX MEDLINE=85216459; PubMed=3858812;
RA Park I., Schaeffer E., Sidoli A., Baralle F.E., Cohen G.N.,
RA Zakin M.M.;
RT "Organization of the human transferrin gene: direct evidence that it
RT originated by gene duplication.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:3149-3153(1985).
```



Db 220 DVAFVHKSTIFENLANKA-----DRDQYELLCLDNTRKPVDBYKDCHEIAQVPSHTV 270  
 QY 282 VVRADTDGG---LIFRLNLEGQRLFSHEGS-SFQMFSSSEAYGKOLLFKDSTSELVPIAT 337  
 Db 271 VAR---SMGKEDLIWELNQAEHFCKDSKEFLFSS-PHG-KDLLFKDSAHGFLKVP 326  
 QY 338 Q-TYEAWLGHYELHAKGLL-----CDPNRLPPYLRWCVLSTPEIKQCGDMAVA.385  
 Db 327 RMDAKMYLGYEYVTAIRNLREGTCPEAPTDECKP-----VKWCALSHHERLKCDEWSV- 379  
 QY 386 FRQRLKPEIQCSAKSPQCHMERIOAEQVDATVLSGEDITYAGKYGLVPAAGEHYAPE 445  
 Db 380 ----NSVGKIECVSAETTEDCTAKIMNGEADMSLDGGFFYTAG-KCGLVPVLAENYKNS 434  
 QY 446 DGSN-----SYVVAVVRDSSHAFTLDELGRKRSCHAGFGSPAGWDVPVGVALIQRGFTIR 500  
 Db 435 DNCEDTPEAGYPAVAVVKSASD-LTWDNLKGKKSCHTAGRTAGWNIPMGLL-----YNK 489  
 QY 501 PKDCDVLTVASEPFPNASCVPVNNPKYPSLSLALCVGDEQGRNKCVCNQSERYGVRGAF 560  
 Db 490 INHC-----RPFDEPFGCAGPSKK---DSSLCKLCMG--SGLNLCEPNNKEGYGYTGAF 540  
 QY 561 RCLVENAGDAVAFVRHTTVDNTNGHNSPEWAAELRSEDYELLCPCNGARAEVSQFAACNLA 620  
 Db 541 RCLVEK-GDVAFAVKHQTVPQNTGGKPPDPWAKNLNEKOYELLCDDGTRKPVBEYANCHLA 599  
 QY 621 QIPPHAVMVRPDTNITVYGLLDKAQDLFGDDHN--KNGFKMFDSSNYHQQDLFFKDATV 678  
 Db 600 RAPNAVVRTRKOEK-CVHKILRQOQLFGNSVNTDCSGNFCFLRSET---KOLLFFRDDIV 655  
 QY 679 RAVPVGSEKTYRGWLGDVYVALEGNMSOQCS 710  
 Db 656 CLAKLHNRNTYKYLGEYVAVGNL--RKCS 685

RESULT 7

TRFE\_HORSE STANDARD; PRT; 706 AA.  
 AC 27425;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serotransferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal binding globulin).  
 GN TF.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93277958; PubMed=8504171;  
 RA Carpenter M.A., Broad T.E.;  
 RT "The cDNA sequence of horse transferrin."  
 RL Biochim. Biophys. Acta 1173:230-232(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Extraembryonic tissue;  
 RA McDowell K.J., Adams M.H., Baker C.B.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Transferrins are iron binding transport proteins which can bind two atoms of ferric iron in association with the binding of an anion, usually bicarbonate. It is responsible for the transport of iron from sites of absorption and heme degradation to those of storage and utilization. Serum transferrin may also have a further role in stimulating cell proliferation.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.  
 CC -!- DOMAIN: Composed of two homologous domains.  
 CC -!- SIMILARITY: Belongs to the transferrin family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M69020; AAA30958.1; -;  
 DR EMBL; U21127; AAA63684.1; -;  
 DR PIR; S33761; S33761.  
 DR HSSP; P02787; LA8E.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 706 SEROTRANSFERRIN.  
 FT REPEAT 20 357 1.  
 FT REPEAT 358 706 2.  
 FT DISULFID 26 64 BY SIMILARITY.  
 FT DISULFID 33 55 BY SIMILARITY.  
 FT DISULFID 134 215 BY SIMILARITY.  
 FT DISULFID 174 190 BY SIMILARITY.  
 FT DISULFID 177 198 BY SIMILARITY.  
 FT DISULFID 187 200 BY SIMILARITY.  
 FT DISULFID 248 262 BY SIMILARITY.  
 FT DISULFID 360 623 BY SIMILARITY.  
 FT DISULFID 366 398 BY SIMILARITY.  
 FT DISULFID 376 389 BY SIMILARITY.  
 FT DISULFID 423 701 BY SIMILARITY.  
 FT DISULFID 441 684 BY SIMILARITY.  
 FT DISULFID 474 550 BY SIMILARITY.  
 FT DISULFID 498 692 BY SIMILARITY.  
 FT DISULFID 508 522 BY SIMILARITY.  
 FT DISULFID 519 533 BY SIMILARITY.  
 FT DISULFID 590 604 BY SIMILARITY.  
 FT DISULFID 642 647 BY SIMILARITY.  
 FT METAL 79 79 BY SIMILARITY.  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 209 209 IRON 1 (BY SIMILARITY).  
 FT METAL 270 270 IRON 1 (BY SIMILARITY).  
 FT METAL 413 413 IRON 2 (BY SIMILARITY).  
 FT METAL 449 449 IRON 2 (BY SIMILARITY).  
 FT METAL 544 544 IRON 2 (BY SIMILARITY).  
 FT METAL 612 612 IRON 2 (BY SIMILARITY).  
 FT BINDING 136 136 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 140 140 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 142 142 CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY).  
 FT BINDING 143 143 CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY).  
 FT BINDING 476 476 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 480 480 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 482 482 CARBONATE 2 (VIA AMIDE NITROGEN) (BY SIMILARITY).  
 FT BINDING 483 483 CARBONATE 2 (VIA AMIDE NITROGEN) (BY SIMILARITY).  
 FT CARBOHYD 515 515 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT SEQUENCE 706 AA; 78094 MW; 1A0FA566C0409D8A CRC64;  
 Query Match 32.1%; Score 1263.5; DB 1; Length 706;  
 Best Local Similarity 41.2%; Pred. No. 1.5e-88;  
 Matches 307; Conservative 113; Mismatches 237; Indels 89; Gaps 30;  
 QY 1 MRGPSALWLLALRTVLGGMEVRWCATSDPEQHKGNMSEAFREAGIOPSL-CVYRGTS 59  
 Db 1 MELATRALACAVLGLCLAEQTVRMCTVSNHVSFKASFRDSMKSVAPPLVACVKRTS 60





QY 540 QGRNCKVGNQRYGYRGAPCLVFNAGDVAFVRHTTVFNTNGHNSPWAELRSBDY 599  
 Db ENKNCMPNSEERYGYTGAFCFLAEKAGDVAFVKDVTVLQNTDCKNSEPWAKDLKQEDF 575  
 QY 600 ELLCPNGARAEVSQFAACNLAIQIPHAVVRPDTHIFTVYGLLDKXQDLFGDDHVNKG-- 657  
 Db ELLCLDGRKPVAAEASCHLAPNHAHVVSQSD-RAQHLKKVFLQDQDFGG----NGPD 630  
 QY 658 ----FKMPDSSNVHQDILLFKDQAVRVPVGVGKTTYRGWLGIDYVAALGKSSQOCSCGAA 713  
 Db 631 CGPKCFLFKSET---KNLLFNNDTECLAEQLQKGTTYEOLGSEYVTSITNL--RRCS--- 682  
 QY 714 APAPGAPLL 722  
 Db 683 ----SSPLL 687

RESULT 9  
 ID TRFL HUMAN STANDARD; PRT; 711 AA.  
 AC P02788; O00756; Q16780; Q16785; Q16786; Q16789; Q96KZ4; Q96KZ5;  
 AC Q9H1Z3;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferroxin A;  
 DE Lactoferroxin B; Lactoferroxin C].  
 GN LTF OR LF.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=90384839; PubMed=2402455;  
 RA Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;  
 RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
 RL Nucleic Acids Res. 18:5288-5288(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Cho Y.Y.;  
 RL Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Connely O.M.;  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Liang Q., Jimenez-Flores R., Richardson T.;  
 RL "Molecular cloning and sequence analysis of human lactoferrin.";  
 RT Submitted (DEC-1991) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RA Wei X., Han J., Rado T.A.;  
 RL "Human neutrophil lactoferrin coding and 5' flanking region DNA  
 sequences.";  
 RT Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Cheng H., Chen X., Huan L.;  
 RT "cDNA cloning and sequence analysis of human lactoferrin.";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Prostate;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [8]  
 RP SEQUENCE OF 3-711 FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=90326549; PubMed=2374734;  
 RA Powell M.J., Ogden J.E.;  
 RT "Nucleotide sequence of human lactoferrin cDNA.";  
 RL Nucleic Acids Res. 18:4013-4013(1990).  
 RN [9]  
 RP SEQUENCE OF 20-711.  
 RX MEDLINE=85076667; PubMed=6510420;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
 RA Legrand D., Spik G., Montreuil J., Jolles P.;  
 RT "Human lactotransferrin: amino acid sequence and structural  
 comparisons with other transferrins.";  
 RL Eur. J. Biochem. 145:659-666(1984).  
 RN [10]  
 RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
 RX MEDLINE=82046817; PubMed=6794640;  
 RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "The present state of the human lactotransferrin sequence. Study and  
 alignment of the cyanogen bromide fragments and characterization of  
 N- and C-terminal domains.";  
 RL Biochim. Biophys. Acta 670:243-254(1981).  
 RN [11]  
 RP SEQUENCE OF 609-711.  
 RX MEDLINE=82262043; PubMed=7049727;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "An 88 amino acid long C-terminal sequence of human  
 lactotransferrin.";  
 RL FEBS Lett. 142:107-110(1982).  
 RN [12]  
 RP SEQUENCE OF 436-711 FROM N.A.  
 RX MEDLINE=88001031; PubMed=3477300;  
 RA Rado T.A., Wei X., Benz E.J. Jr.;  
 RT "Isolation of lactoferrin cDNA from a human myeloid library and  
 expression of mRNA during normal and leukemic myelopoiesis.";  
 RL Blood 70:989-993(1987).  
 RN [13]  
 RP SEQUENCE OF 237-711 FROM N.A.  
 RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
 RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
 RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
 RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
 RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
 RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
 RA Sagripanti J.L.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
 RX MEDLINE=90064528; PubMed=2585506;  
 RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
 RT "Structure of human lactoferrin: crystallographic structure analysis  
 and refinement at 2.8-A resolution.";



SEQUENCE FROM N.A.  
 RC TISSUE-Mammary gland;  
 RA Lee T., Yu S., Kim S., Lee K., Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Mammary gland;  
 RX MEDLINE=94380047; PubMed=8093048;  
 RA le Provost F., Nocard M., Guerin G., Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 relevant locus to bovine U12 syntenic group.";  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).  
 CC !- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE.  
 CC !- SUBUNIT: Monomer (BY SIMILARITY).  
 CC !- DOMAIN: Composed of two homologous domains.  
 CC !- SIMILARITY: Belongs to the transferrin family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; U53857; AAA97958.1; -;  
 DR EMBL; X78902; CAA55517.1; -;  
 DR HSSP; 077698; 1CB2.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TF\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN 1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN 2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN 3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.  
 FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 2 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 142 142 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 SIMILARITY).  
 FT BINDING 143 143 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 SIMILARITY).

FT BINDING 478 478 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 482 482 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 484 484 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 SIMILARITY).  
 FT BINDING 485 485 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 56 56 I -> V (IN REF. 2).  
 FT CONFLICT 88 88 L -> R (IN REF. 2).  
 FT CONFLICT 124 124 Q -> K (IN REF. 2).  
 FT CONFLICT 154 154 F -> P (IN REF. 2).  
 FT CONFLICT 304 304 S -> R (IN REF. 2).  
 FT CONFLICT 414 414 D -> G (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77358 MW; F2EDA3C83539960D CRC64;  
 Query Match 32.0%; Score 1258; DB 1; Length 708;  
 Best Local Similarity 40.8%; Pred. No. 3.9e-88;  
 Matches 305; Conservative 110; Mismatches 243; Indels 90; Gaps 30;  
 QY 6 GALMLLLALRTVLGGMEVRWCATSDPEQHKCCNMSEAFREAGIQPSLLCVRGTSADHCVQ 65  
 DB 12 GALGLCLAAPR-----KNVRWCAISLPWCKYQWQRMKLG-APSIICIRTSALECIR 66  
 QY 66 LIAQEAQDAITLDGAIYEAGKE-HGLKPVVGEVVDQEVG--TSYYAVAVVRSSHVITD 122  
 DB 67 AIAGKNADAVTLDSGMVFEGALDPKYLRFAAEIYGTESKSPQTHYYAVAVVKGSNFOLD 126  
 QY 123 TLKGVKSCHTGINRTVGNVVPVGYL-----VESGRLSVMGCDVLKVDYDGGSCVPGA 176  
 DB 127 QLQGQKSCMGLGRSAGWNI PVGILRPLSWTESAE-----PLOGAVARFFSACVPCV 180  
 QY 177 GETSYSESLCRICRGDSSGEGVCDKSPLEYYDYSAGFRCLEAGDGVAFVKHSTVLENT 236  
 DB 181 DGKAY-PNLQCLCKG--VGENKCASSQEPFYGSYGFKCLQDAGDGVAFVKETTVFEN- 236  
 QY 237 DGKTLPSWGQALLSQDFELLCRDGRADVTWRQCHLARVPAHVVVRA-DTDGGLIFRL 295  
 DB 237 ---LPERAD---RDQYELLCLNNTRAPVDAFKECHLAQVPSHAVVARSVDGKENVL 289  
 QY 296 LNEGQRLP-SHEGSSFOMFSEAYGOKDLLFKDSTSELVPIATOTYEA-WLGEHYLHAMK 353  
 DB 290 LRKAQEKFGNKSQSFQLFGSPE-GRDELFPDSALGFVRIPSKVDSALYLSRYLTALK 348  
 QY 354 GILCDPNRPL---PYLRMCVLSTPEIQKCGDMAVAFRRQRRLKPEIQCVSAKSPQHCME 410  
 DB 349 NLRETAEBELKARCTRVVWCAVGPESQCKQWS-----EQSQNVTCATASTDDCIALV 403  
 QY 411 QAEQVDVAVTLGSDIYTAGKYGVLVPAAGEH-----YAPED-----SSNSYVVAVVRDS 461  
 DB 404 LKGEADALSLDGGYIYTAG-KCGLVPVMAENRKSSKSYSLDCVLRPTTEGYLAVAVVVK-A 461  
 QY 462 SHAFITDELGRKRSCHAGFGSPAGWDVPVGGALI-ORGFIRPKCDVLTAVSFFENASCVP 520  
 DB 462 NEGLTWNLSLKGKKSCHTAVDRTAGWNI PMGLIANTGT-----SC-----AFDEFFSQSCAP 512  
 QY 521 VNNPKNYPSSICALCVGDEQGRNKCVCNSQRYGYRGAFRCLEVENAGDGVAFVRHTTVPD 580  
 DB 513 GADPK---SSJLCALCAGDQGLDKCVPNSKEKYGYTGAFRCLEADVEDGVAFVKNDTWE 569  
 QY 581 NTNGHNSFPWAAELRSEDYELLCPNGARAEVYSQFACNLAIQIPPHAVVVRPDTNIFTVYG 640  
 DB 570 NTNGESSADWAKNLNRDFRLCLDGTGTPVTEAQSCLAVAPNHAVVRSRSDRAAHEVO 629  
 QY 641 LLDKAQLDQDDHKNKG-----FKWFDSSNVHGDLLFKDATVRAVPVGEKTTVRMLG 694  
 DB 630 LLHQ-QALFG-----KNGKNCPDQPCLFKSET---KNLLFNQNTCLAKLGRPTVEKYL 681  
 QY 695 LDYVAALGEGSSQCSGAAAPAPGAPLL 722  
 FT BINDING 478 478 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 482 482 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 484 484 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 SIMILARITY).  
 FT BINDING 485 485 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 56 56 I -> V (IN REF. 2).  
 FT CONFLICT 88 88 L -> R (IN REF. 2).  
 FT CONFLICT 124 124 Q -> K (IN REF. 2).  
 FT CONFLICT 154 154 F -> P (IN REF. 2).  
 FT CONFLICT 304 304 S -> R (IN REF. 2).  
 FT CONFLICT 414 414 D -> G (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77358 MW; F2EDA3C83539960D CRC64;





CC -1- DOMAIN: Composed of two homologous domains.  
CC -1- SIMILARITY: Belongs to the transferrin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AJ131674; CAB53387.1; -  
CC EMBL; AF165879; AAF82241.1; -  
CC PDB; 1DTZ; 20-JUN-01.  
CC InterPro; IPR001156; Transferrin.  
CC Pfam; PF00405; transferrin; 2.  
CC PRINTS; PR00422; TRANSFERRIN.  
CC SMART; SM00094; TR\_FER; 2.  
CC PROSITE; PS00205; TRANSFERRIN 1; 2.  
CC PROSITE; PS00206; TRANSFERRIN 2; 2.  
CC PROSITE; PS00207; TRANSFERRIN 3; 2.  
CC Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
KW Signal; 3D-structure  
FT SIGNAL 1 19 BY SIMILARITY.  
FT CHAIN 20 708 LACTOTRANSFERRIN.  
FT REPEAT 20 363 1.  
FT REPEAT 364 708 2.  
FT DISULFID 28 64 BY SIMILARITY.  
FT DISULFID 38 55 BY SIMILARITY.  
FT DISULFID 134 217 BY SIMILARITY.  
FT DISULFID 176 192 BY SIMILARITY.  
FT DISULFID 189 200 BY SIMILARITY.  
FT DISULFID 250 264 BY SIMILARITY.  
FT DISULFID 367 399 BY SIMILARITY.  
FT DISULFID 377 390 BY SIMILARITY.  
FT DISULFID 424 703 BY SIMILARITY.  
FT DISULFID 444 666 BY SIMILARITY.  
FT DISULFID 476 551 BY SIMILARITY.  
FT DISULFID 500 694 BY SIMILARITY.  
FT DISULFID 510 524 BY SIMILARITY.  
FT DISULFID 521 534 BY SIMILARITY.  
FT DISULFID 592 606 BY SIMILARITY.  
FT DISULFID 644 649 BY SIMILARITY.  
FT METAL 79 79 IRON 1 (BY SIMILARITY).  
FT METAL 111 111 IRON 1 (BY SIMILARITY).  
FT METAL 211 211 IRON 1 (BY SIMILARITY).  
FT METAL 272 272 IRON 1 (BY SIMILARITY).  
FT METAL 414 414 IRON 2 (BY SIMILARITY).  
FT METAL 452 452 IRON 2 (BY SIMILARITY).  
FT METAL 545 545 IRON 2 (BY SIMILARITY).  
FT METAL 614 614 IRON 2 (BY SIMILARITY).  
FT BINDING 136 136 CARBONATE 1 (BY SIMILARITY).  
FT BINDING 140 140 CARBONATE 1 (BY SIMILARITY).  
FT BINDING 142 142 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
FT SIMILARITY).  
FT BINDING 143 143 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
FT SIMILARITY).  
FT BINDING 478 478 CARBONATE 2 (BY SIMILARITY).  
FT BINDING 482 482 CARBONATE 2 (BY SIMILARITY).  
FT BINDING 484 484 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
FT SIMILARITY).  
FT BINDING 485 485 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
FT SIMILARITY).  
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 261 261 F -> S (IN REF. 2).  
FT CONFLICT 304 304 G -> A (IN REF. 2).  
FT CONFLICT 330 330 S -> P (IN REF. 2).  
FT CONFLICT 392 392 LLS -> PLF (IN REF. 2).  
FT CONFLICT 492 494 L -> F (IN REF. 2).  
FT CONFLICT 506 506 A -> P (IN REF. 2).  
FT CONFLICT 609 609

FT CONFLICT 642 642 R -> Q (IN REF. 2).  
SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;  
Query Match 31.7%; Score 1249; DB 1; Length 708;  
Best Local Similarity 40.0%; Pred. No. 1.9e-87;  
Matches 297; Conservative 122; Mismatches 245; Indels 78; Gaps 27;  
QY 6 GALWLLALRTVLGMEVRWCATSDPEQHKKCNMSEAFREAGIQPSLLCVRGTSADHCQV 65  
DB 12 GALGLCLA---ASKSVRWCTTSPAESSKCAQWRMKKVR-GPSVTCVKKTSRECIQ 66  
QY 66 LTAQGEADAITLDGGAIEAGKE-HGLKPVCEVVDQEVG--TSYYAVAVRRSSHTVD 122  
DB 67 AISTEKADAVTLDGLVWDAGLDPPYKLRFAAEVYGTENNPPQTHYAVAIKKGTNFQLN 126  
QY 123 TLKGVKSCHTGINRTVGNVPVGLVESGRSLVSMGCDVLKAVSDVFGGSCVPCGAGTSYS 182  
DB 127 QLOGLKSCHTGLGRSAGNIPMGLLRPFLDWTGPPEPQKAVAKAFSASCVCVDCGKEY- 185  
QY 183 ESLCRLCRGDSGEGVCDKSPLEYYDYSAGFCLAEAGDGVAFVKHSTVLENTDGTLP 242  
DB 186 PNLCLQAG--TGENKCACSQEPYFGYSAGFKCLQDGAGDVAFAVKDSTVFESLPAKA-- 241  
QY 243 SWGQALLSQD-FELICRDSRADVTWROCHLARUPAHAVVRA-DTDCGLIFRLNBEQ 300  
DB 242 -----DRDQYELLCPNNTKRPVDAFQECHLARVPASHAVARSVNGKEDLIWKLVAQ 294  
QY 301 RLFSH-EGSSFQMFSEAVGKDLDFKDDSTSLVPIATOTYEA-WLGHVYHAKMGL--- 355  
DB 295 EXFGKPGSGFQFGSPA-GQNDLLFKDSALGLLRISKIDSGLYLGSNYITAIRGLRET 353  
QY 356 LCDPNRLPYLRWCVLSTPEIQKCGDMAVAFRRQRKLPKIQCVSAKSQPHCMERIOAEQV 415  
DB 354 AAEVELRRQAQVVCVAVGSDEQLKCEWS-----RQSNQSVVCATASTTDCIALVLKGEA 408  
QY 416 DAVTSLGEDIYTAGKYGLVPAAGE-HYAPESS-----NSYVAVAVRRDSSHAFT 466  
DB 409 DALSLDGGYIYTAG-KCGLVPVLAESQSPSSGLDCVHRPVKGYLAVALAVVRK-ANDKIT 466  
QY 467 LDELGRKRSCHAGFGSPAGWDVPVGLIQRGFIKPKCDVLTAVSFFNASCPVNVNPKN 526  
DB 467 WNSLRGKSKCHTAVDRTAGNIPMGLL-----SKNTD-SCRFDEFLSQSCAPGSDPR- 517  
QY 527 YPSSLCALCVGDEQGRNKCVCNSQRYGYRGAFCLVENAGDVAFRHTTTFDNTNGHN 586  
DB 518 --SKLCAACAGNEQNKVCVNSERYGYTCFACLAENVGDVAFVKDVTVLNDTNGKN 575  
QY 587 SEPMAELRSEDYELLCPNNGARAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKQ 646  
DB 576 TEQWAKDLKLDGDFELLCNLGTRKPVTEAESCHLAVAPNHAVVSRID-KVAHLEQVLLRQ 634  
QY 647 DLFGDDHKNKG-----FKMFDSSNVHGQDLFKDATVRAVPVGEKTYRGMGLDYVAA 700  
DB 635 AHFG-----RNGRDCPKFKCLFQSKT---KNLLFNONTCLAKLQGGKTYEELGPOYVTA 687  
QY 701 LEGMSSQCSGAAAPAPGAPLL 722  
DB 688 IAKL--RRCS-----TSPLL 700  
RESULT 13  
TRFE ONCKI STANDARD; PRT; 687 AA.  
ID TRFE ONCKI AC P79815;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serotransferrin precursor.  
OS Oncorhynchus kisutch (Coho salmon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8019;



RX MEDLINE=91274362; PubMed=2054387;  
RA Banfield D.K., Chow B.K.-C., Funk W.D., Robertson K.A., Umelas T.M.,  
RA Woodworth R.C., Macgillivray R.T.A.;  
RT "The nucleotide sequence of rabbit liver transferrin cDNA.";  
RL Biochim. Biophys. Acta 1089:262-265(1991).  
RN [2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=New Zealand white;  
RA Ghareeb B.A.A., Thepot D., Puisant C., Cajero-Juarez M.,  
RA Houdebine L.M.;  
RT "Cloning and structural organisation of the rabbit transferrin encoding  
RT gene";  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE OF 20-51.  
RP MEDLINE=88209278; PubMed=3365331;  
RA Godovac-Zimmermann J.;  
RT "Isolation, characterization and N-terminal amino-acid sequence of  
RT rabbit transferrin.";  
RL Biol. Chem. Hoppe-Seyler 369:93-96(1988).  
RN [4]  
RN SEQUENCE OF 483-545.  
RP MEDLINE=89005676; PubMed=3169252;  
RA Evans R.W., Aitken A., Patel K.J.;  
RT "Evidence for a single glycan moiety in rabbit serum transferrin and  
RT location of the glycan within the polypeptide chain.";  
RL FEBS Lett. 238:39-42(1988).  
RN [5]  
RN X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).  
RC TISSUE=Plasma;  
RX MEDLINE=89026775; PubMed=3179277;  
RA Bailey S., Evans R.W., Garratt R.C., Gorinsky B., Hasnain S.,  
RA Horsburgh C., Jhoti H., Lindley P.F., Mydin A., Sarra R., Watson J.L.;  
RT "Molecular structure of serum transferrin at 3.3-A resolution.";  
RL Biochemistry 27:5804-5812(1988).  
RN [6]  
RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
RA Sarra R., Garratt R.C., Gorinsky B., Jhoti H., Lindley P.F.;  
RT "High-resolution X-ray studies on rabbit serum transferrin:  
RT preliminary structure analysis of the N-terminal half-molecule at  
RT 2.3-A resolution.";  
RL Acta Crystallogr. B 46:763-771(1990).  
CC -1- FUNCTION: Transferrins are iron binding transport proteins which  
CC can bind two atoms of ferric iron in association with the binding  
CC of an anion, usually bicarbonate. It is responsible for the  
CC transport of iron from sites of absorption and heme degradation to  
CC those of storage and utilization. Serum transferrin may also have  
CC a further role in stimulating cell proliferation.  
CC -1- SUBUNIT: Monomer.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.  
CC -1- DOMAIN: Composed of two homologous domains.  
CC -1- SIMILARITY: Belongs to the transferrin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X58533; CAA41424.1; --  
DR EMBL; AF031625; AAB94136.1; --  
DR EMBL; AF031611; AAB94136.1; JOINED.  
DR EMBL; AF031612; AAB94136.1; JOINED.  
DR EMBL; AF031613; AAB94136.1; JOINED.  
DR EMBL; AF031614; AAB94136.1; JOINED.  
DR EMBL; AF031615; AAB94136.1; JOINED.  
DR EMBL; AF031616; AAB94136.1; JOINED.  
DR EMBL; AF031617; AAB94136.1; JOINED.  
DR EMBL; AF031618; AAB94136.1; JOINED.  
DR EMBL; AF031619; AAB94136.1; JOINED.  
DR EMBL; AF031620; AAB94136.1; JOINED.  
DR EMBL; AF031621; AAB94136.1; JOINED.  
DR EMBL; AF031622; AAB94136.1; JOINED.  
DR EMBL; AF031623; AAB94136.1; JOINED.  
DR EMBL; AF031624; AAB94136.1; JOINED.  
DR PDB; 1TFD; 15-APR-93.  
DR PDB; 1JNF; 09-JAN-02.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 2.  
DR PRINTS; PR00422; TRANSFERRIN.  
DR SMART; SM00094; TR\_FER; 2.  
DR PROSITE; PS00205; TRANSFERRIN\_1; 1.  
DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
KW Signal; 3D-structure.  
FT SIGNAL 1 19  
FT CHAIN 20 695 SEROTRANSFERRIN.  
FT REPEAT 20 355 1.  
FT REPEAT 356 695 2.  
FT DISULFID 28 67  
FT DISULFID 38 58  
FT DISULFID 137 213  
FT DISULFID 156 350  
FT DISULFID 177 193  
FT DISULFID 180 196  
FT DISULFID 190 198  
FT DISULFID 246 260  
FT DISULFID 358 612  
FT DISULFID 364 396  
FT DISULFID 374 387  
FT DISULFID 421 690  
FT DISULFID 436 653  
FT DISULFID 468 539  
FT DISULFID 492 681  
FT DISULFID 502 516  
FT DISULFID 513 522  
FT DISULFID 579 593  
FT DISULFID 631 636  
FT CARBOHYD 509 509 N-LINKED (GLCNAC. . .).  
FT METAL 82 82 IRON 1.  
FT METAL 114 114 IRON 1.  
FT METAL 207 207 IRON 1.  
FT METAL 268 268 IRON 1.  
FT METAL 411 411 IRON 2.  
FT METAL 444 444 IRON 2.  
FT METAL 533 533 IRON 2.  
FT METAL 601 601 IRON 2.  
FT BINDING 139 139 CARBONATE 1.  
FT BINDING 143 143 CARBONATE 1.  
FT BINDING 145 145 CARBONATE 1 (VIA AMIDE NITROGEN).  
FT BINDING 146 146 CARBONATE 1 (VIA AMIDE NITROGEN).  
FT BINDING 470 470 CARBONATE 2.  
FT BINDING 474 474 CARBONATE 2.  
FT BINDING 476 476 CARBONATE 2 (VIA AMIDE NITROGEN).  
FT BINDING 477 477 CARBONATE 2 (VIA AMIDE NITROGEN).  
FT VARIANT 517 517 V -> I.  
FT CONFLICT 7 7 MISSING (IN REF. 1).  
FT CONFLICT 47 47 K -> S (IN REF. 3).  
FT CONFLICT 50 50 P -> Y (IN REF. 3).  
FT STRAND 27 29  
FT HELIX 32 45  
FT TURN 46 48  
FT TURN 59 60  
FT HELIX 64 72  
FT TURN 73 74  
FT STRAND 78 81  
FT HELIX 83 90  
FT TURN 91 94  
FT STRAND 96 102  
FT STRAND 114 121  
FT TURN 122 123  
FT TURN 128 129

```

FT TURN 132 133
FT STRAND 136 137
FT TURN 141 142
FT TURN 144 147
FT TURN 148 154
FT TURN 155 157
FT TURN 165 170
FT TURN 171 172
FT TURN 176 177
FT TURN 179 180
FT TURN 190 190
FT TURN 194 195
FT TURN 207 214
FT TURN 215 217
FT TURN 221 225
FT TURN 226 227
FT TURN 228 232
FT TURN 236 239
FT TURN 240 241
FT TURN 242 246
FT TURN 247 249
FT TURN 250 253
FT TURN 254 255
FT TURN 257 259
FT TURN 263 265
FT TURN 269 273
FT TURN 279 288
FT TURN 289 291
FT TURN 303 304
FT TURN 306 306
FT TURN 307 308
FT TURN 309 309
FT TURN 316 317
FT TURN 321 322
FT STRAND 321 322
SQ SEQUENCE 695 AA; 76670 MW; DB12F34D87AE9D55 CRC64;

Query Match
Best Local Similarity 41.3%; Score 1240; DB 1; Length 695;
Matches 308; Conservative 108; Mismatches 231; Indels 98; Gaps 34;

QY 1 MRGPGALWLLALRTVLGME--VRWCATSDPEQHKCGN-----MSEAFREAGIQPSLAC 54
DB 1 MRLAAGALLACAALGLCLAVTEKTVRWCAVNDHEASKCANFRDSMKKVLPEDG--PRIIC 58
QY 55 VRGTSADHCVOIARQEAADITLDGAIYEAG-KEHGLKPVVGEVYDQEVG--TSYYAVA 111
DB 59 VKKASYLDCIKIAAHEADAVTLDAGLVHEAGLTPNNLKPVVVAEFGSKENPKTFYYA 118
QY 112 VVRRSHVITDITLKGKVSCHTGINRTVGWNPVGVYLVESGRLSVMGCD-----VLKAV 164
DB 119 LVKGSNFQINELQKKKCHTGLGRSAGWNIFIGLLY-----CDLPPEPRKPLEKAV 169
QY 165 SDYFGSCVPGAGETSYSESLCRLCRGDSGEGVCDKSPLERYDYSGAFRCFLAEGAGDV 224
DB 170 ASFFSGSCVPCADGADFPQ-LQLCPG-----CGCSSVQVPYFGYSGAFKCLKDGLGDV 221
QY 225 AFVKHSTVLNTDGTLPSPGQALLSQDFELLCRDGRADVTWEQCHLARVPAAVVR 284
DB 222 AFVQKQETIFEN-----LPSKDE---RDQYELLCLDNTKRPVDEYEQCHLARVPASHAVAR 273
QY 285 A-DTGGGLIFRLNQRQLFSHEGS-SFQMFSSSEAYGOKDLLFKDSTSELVPIATQ-TYE 341
DB 274 SVDGKEDLLIWEILNQAEHFQKDGDFQFSS-PHG-KVLLFKDSAYGFFKVPFRMDAN 331
QY 342 AWLGHEYLHAMKGL-----LCDPNRLP---PYLRWCVLSTPEIQKCGDMAVAFRRQLKPEI 395
DB 332 LVLGYEYVAVRNLRREGIC-PDPLQDECKAVKWCALSHERLKCDEWSVTSGGL-----I 385
QY 396 QCVSAKSPQHCERIOAEQVDATLSGEDIIYTAGKYGLVPAAGEHY-----APESS 448
DB 386 ECESAEETPDCTAKINMGNEADAMSLDGGVYTAG-QCGLVPVLAENYESTDCKKAPES-- 442
QY 449 NSYVVAVVVRDSSHAFTLDELGRKRSCHAGFGSPAGMDVPVVGALIQRFIRPKDCDVL 508

```

```

DB 443 -CYLSVAVVKK-SNPDINWNNLEGKKSCHTAVDRTAGNIPMGLL-----YNRINHC--- 492
QY 509 AVSEFFNASCVPVNNPKNYPSSLCALCVGDGQGRNKCVCNQSQERYGYGVRGAPRCLVENAG 568
DB 493 RFDEFROGCAP-GSQKN--SSLCLCVGP-----SVCAPNNREGYGYGTGAPRCLVEK-G 544
QY 569 DVAFVYRHTTVPDNTGNHNSPAAELRSEDYELLCPCNGARAEVSQFAACNLAIQIPPHAVM 628
DB 545 DVAFVYKQTVLQNTGGRNSPWAOKLKEEDFELLCDDGTRKPVSEAHNCHLAKANHAVV 604
QY 629 VRPDNTNIFTVYGLLD---KAODLFGDDHNNKNGKFMFDSNHYHGQDLLFKDATTVAVPVGE 685
DB 605 SRKDAKACVCKQLDLQVEYGNVTADCSK--FCNFHSKT-----KOLLFRDDTKCLVDLRG 659
QY 686 KTYRGWLGLDYVAALGEMSSQOCS 710
DB 660 KNTYEKYLGADEYKAVSNL--RKCS 682

RESULT 15
TRFE PAROL
ID TRFE PAROL STANDARD; PRT; 685 AA.
AC O93429;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serotransferrin precursor.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthysidae; Paralicthys.
OC NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim Y., Lee J., Hong Y., Hiroo I., Aoki T.;
RT "Molecular cloning and sequence analysis of transferrin cDNA from
RT Japanese flounder Paralicthys olivaceus.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: Composed of two homologous domains.
CC -!- SIMILARITY: Belongs to the transferrin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D88801; BAA28944.1; -.
DR HSP; P56410; IAOV.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 1.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW Signal.
FT SIGNAL 1 16 BY SIMILARITY.
FT CHAIN 17 685 SEROTRANSFERRIN.
FT DISULFID 26 48 BY SIMILARITY.
FT DISULFID 125 206 BY SIMILARITY.
FT DISULFID 170 184 BY SIMILARITY.
FT DISULFID 234 248 BY SIMILARITY.

```

|  |          |   |     |   |
|--|----------|---|-----|---|
| FT   | DISULFID | 343   | 379 | BY SIMILARITY.                                    |
| FT   | DISULFID | 353   | 370 | BY SIMILARITY.                                    |
| FT   | DISULFID | 404   | 678 | BY SIMILARITY.                                    |
| FT   | DISULFID | 419   | 639 | BY SIMILARITY.                                    |
| FT   | DISULFID | 451   | 526 | BY SIMILARITY.                                    |
| FT   | DISULFID | 475   | 667 | BY SIMILARITY.                                    |
| FT   | DISULFID | 485   | 499 | BY SIMILARITY.                                    |
| FT   | DISULFID | 496   | 509 | BY SIMILARITY.                                    |
| FT   | DISULFID | 566   | 580 | BY SIMILARITY.                                    |
| FT   | METAL    | 72  | 72  | IRON 1 (BY SIMILARITY).                           |
| FT   | METAL    | 102   | 102 | IRON 1 (BY SIMILARITY).                           |
| FT   | METAL    | 200   | 200 | IRON 1 (BY SIMILARITY).                           |
| FT   | METAL    | 256   | 256 | IRON 1 (BY SIMILARITY).                           |
| FT   | METAL    | 394   | 394 | IRON 2 (BY SIMILARITY).                           |
| FT   | METAL    | 428   | 428 | IRON 2 (BY SIMILARITY).                           |
| FT   | METAL    | 520   | 520 | IRON 2 (BY SIMILARITY).                           |
| FT   | METAL    | 588   | 588 | IRON 2 (BY SIMILARITY).                           |
| FT   | BINDING  | 127   | 127 | CARBONATE 1 (BY SIMILARITY).                      |
| FT   | BINDING  | 131   | 131 | CARBONATE 1 (BY SIMILARITY).                      |
| FT   | BINDING  | 133   | 133 | CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY). |
| FT   | BINDING  | 134   | 134 | CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY). |
| FT   | BINDING  | 453   | 453 | CARBONATE 2 (BY SIMILARITY).                      |
| FT   | BINDING  | 457   | 457 | CARBONATE 2 (BY SIMILARITY).                      |
| FT   | BINDING  | 459   | 459 | CARBONATE 2 (VIA AMIDE NITROGEN) (BY SIMILARITY). |
| FT   | BINDING  | 460   | 460 | CARBONATE 2 (VIA AMIDE NITROGEN) (BY SIMILARITY). |
| FT   | CARBOHYD | 476   | 476 | N-LINKED (GLCNAC...) (POTENTIAL).                 |
| FT   | SEQUENCE | 685 AA; 74654 MW; 5A6F622AD7D5B662 CRC64;                     |     |   |
| Query Match 31.4%; Score 1236.5; DB 1; Length 685;                 |          |   |     |   |
| Best Local Similarity 38.8%; Pred. No. 1.6e-86;                    |          |   |     |   |
| Matches 277; Conservative 118; Mismatches 242; Indels 77; Gaps 24; |          |   |     |   |
| Qy   | 10       | LLALRTVLGME-VRWCATSDPEQHKCGNMSAFREAGIQPSLLCVRGTSADHCVOLIA     | 68  |   |
| Db   | 9        | LLGLATIASCIDTVKNCVTSTKENLKTALAAA-----APVFCVARASITDCLTAIK      | 62  |   |
| Qy   | 69       | AQADATLGGAIYAG-KEHGLKPVGEVYQEVGTSYVAVVRRSHVTDITLKG            | 127 |   |
| Db   | 63       | AGEADATLGGGEIYTAGLDEYKLPHTIAEQYGTSTDTCTCYAVAVAKKNTGFLHQLMGK   | 122 |   |
| Qy   | 128      | KSCHTGINRTVGNVPVGYLVESGRSLVMGCDVLK---AVSDYFGGSCVPAGAGETSSES   | 184 |   |
| Db   | 123      | KSCHTGVGKSAGWNPITGILLSDMDFIKWKGSDDKLEEVVGEFFHSCAPGATD---SAN   | 179 |   |
| Qy   | 185      | LCRLCRGDSGEGVCDK-SPLERYDYSGAFRCLEAGAGDVAFYKHSVLENTDGTPLPS     | 243 |   |
| Db   | 180      | LCKLCIGD-----CSKSSSETPYNYHGAFCQCLKDGKGDVAFVKHLLVPEEB-----     | 226 |   |
| Qy   | 244      | WQALLSQDELLCRDGSRADVTWEQCHLARPAPAHAVVVRADTD---GGLI FRLNREGQR  | 301 |   |
| Db   | 227      | -----KNDYELLCCKONTXRPIDQFENDLAKVPASHAVTRKONEELAQFIWQSLSVK-    | 279 |   |
| Qy   | 302      | LFSHEGSSFOMFSEAYGQKDLFPKDSLSELPATOT-YEAWLHGEYHAKMGL--LCD      | 358 |   |
| Db   | 280      | -----NENLPSSTPYGKNLNFKDSSTTLVQLPLNVDTMYLGPHYLESVALKIUNI       | 332 |   |
| Qy   | 359      | PNRLPPYLRCVLSTPEIQCGDMAVAFRRQRLKPEIQCVSAKSPQHCHMERIOAEQVAV    | 418 |   |
| Db   | 333      | PSITSDAMKWCAGVRSBDCDSWSVASLVQD-GTTIDCIKGNVTDDCLKKIMHKEADAM    | 391 |   |
| Qy   | 419      | TLSGEDIYTAGKYGLVPAAGEHY-----APEDSSNSYVAVVVRDSSHAFTLDELRG      | 472 |   |
| Db   | 392      | AVDGGQYITAG-KCGLVPAMVEQYDQGCSCAP-GAARLYYAVAVIKKS--GVTWENLRN   | 447 |   |
| Qy   | 473      | KRSCHAGFGSPAGWDVPVVGALIQGFIRPKDCDVLTAVERSEFFNASCVPVNPNKYPSSLC | 532 |   |
| Db   | 448      | KRSCHTIGIRNAGWNTPMGLIYE-----QTKNCN-----FSAFFSSSCAPGADPS---    | 496 |   |
| Qy   | 533      | ALCVGDEQGRNKCVCNQSERYGYRGAFRCLEVENAGDVAFVRHTTVTDNTNGHNSBPAA   | 592 |   |

Search completed: May 14, 2004, 09:42:20  
Job time : 20 secs

|    |     |  |     |
|----|-----|--|-----|
| Db | 497 | AQCAGNAESINKCKASNEERYAYAGAFRCLEAGKGDVAFVKHSIVKENTDQGPPE-WAK  | 555 |
| Qy | 593 | ELRSEDEYELLCPNGARAEVSQFAACNLAQIPPHAVVVRPD--TNIFTVYGLLQKQDLFG | 650 |
| Db | 556 | AFLSNDYELICPSKGPVSVENFMSCNLAKVNAHAVVTRPEIRTKVVT---FLNNQOSHFG | 612 |
| Qy | 651 | DDHNKNGPKMFDSNHYHQDQLLFKDATVRAVPVGEKTTYRGWLGLDYVAALEGM       | 704 |
| Db | 613 | NSASEESFKMFTSPD--GENLLFKYSTKCLQEI PAHLDYKGLGQYMTVMSSL        | 664 |

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 09:40:37 ; Search time 23 Seconds  
(without alignments)  
1656.520 Million cell updates/sec

Title: US-10-049-957-4  
Perfect score: 3936  
Sequence: 1 MRGPSALWLLALRTVLGG.....APLPLLLPALAARLLPPAL 738

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description      |
|------------|--------|-------------|--------|-------|------------------|
| 1          | 3933   | 99.9        | 738    | 6     | 5262177-2        |
| 2          | 3840   | 97.6        | 719    | 2     | US-08-520-933-3  |
| 3          | 3840   | 97.6        | 719    | 4     | US-09-285-040-3  |
| 4          | 3792   | 96.3        | 717    | 6     | 5262177-5        |
| 5          | 1850   | 47.0        | 502    | 2     | US-08-459-818-19 |
| 6          | 1850   | 47.0        | 502    | 2     | US-08-989-666-19 |
| 7          | 1850   | 47.0        | 502    | 2     | US-08-465-078-19 |
| 8          | 1850   | 47.0        | 502    | 2     | US-08-725-776-19 |
| 9          | 1850   | 47.0        | 502    | 2     | US-08-488-062-19 |
| 10         | 1269.5 | 32.3        | 698    | 2     | US-08-175-158A-2 |
| 11         | 1268.5 | 32.2        | 1074   | 2     | US-08-470-058-2  |
| 12         | 1268.5 | 32.2        | 1074   | 3     | US-09-037-188-2  |
| 13         | 1268.5 | 32.2        | 1074   | 3     | US-09-285-310-2  |
| 14         | 1263.5 | 32.1        | 1410   | 2     | US-08-470-058-4  |
| 15         | 1263.5 | 32.1        | 1410   | 3     | US-09-037-188-4  |
| 16         | 1263.5 | 32.1        | 1410   | 3     | US-09-285-310-4  |
| 17         | 1261   | 32.0        | 705    | 2     | US-08-655-640-2  |
| 18         | 1260.5 | 32.0        | 708    | 2     | US-08-555-640-4  |
| 19         | 1255.5 | 31.9        | 711    | 1     | US-08-145-681-2  |
| 20         | 1255.5 | 31.9        | 711    | 1     | US-08-250-308-2  |
| 21         | 1255.5 | 31.9        | 711    | 1     | US-08-453-703-2  |
| 22         | 1255.5 | 31.9        | 711    | 1     | US-08-456-106-2  |
| 23         | 1255.5 | 31.9        | 711    | 3     | US-08-456-108-2  |
| 24         | 1255.5 | 31.9        | 711    | 3     | US-09-265-577-2  |
| 25         | 1255.5 | 31.9        | 711    | 4     | US-09-633-739-2  |
| 26         | 1255.5 | 31.9        | 711    | 5     | PCT-US93-03614-2 |
| 27         | 1252   | 31.8        | 694    | 3     | US-08-724-586-2  |

ALIGNMENTS

RESULT 1  
5262177-2

; Patent No. 5262177  
; APPLICANT: BROWN, J OSEPH P.; ESTIN, CHARLES D.; PLOWMAN, GREGORY  
; D.; HELSTROM, KARL E.; ROSE, TIMOTHY M.; HELSTROM, INGERGERD;  
; PURCHIO, ANTHONY F.; HU, SHIU-LOK; PENNATHUR, SRIDHAR  
; TITLE OF INVENTION: RECOMBINANT VIRUSES ENCODING THE HUMAN  
; MELANOMA-ASSOCIATED ANTIGEN  
; NUMBER OF SEQUENCES: 6  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/7230  
; FILING DATE: 27-JAN-1987  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 827,313  
; FILING DATE: 07-FEB-1986  
; SEQ ID NO: 2:  
; LENGTH: 738  
5262177-2

Query Match 99.9%; Score 3933; DB 6; Length 738;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 737; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

|    |     |  |     |
|----|-----|--|-----|
| Qy | 1   | MRGPSALWLLALRTVLGGMEVRCATSDPEQHKCGNMSEAPREAGIQPSLLCVRGTS     | 60  |
| Db | 1   | MRGPSALWLLALRTVLGGMEVRCATSDPEQHKCGNMSEAPREAGIQPSLLCVRGTS     | 60  |
| Qy | 61  | DHCVQLIAAQEADAITLDGAIYEAQKEHGLKPVGVEVDQEVGTSTYAVAVVRRSSHVT   | 120 |
| Db | 61  | DHCVQLIAAQEADAITLDGAIYEAQKEHGLKPVGVEVDQEVGTSTYAVAVVRRSSHVT   | 120 |
| Qy | 121 | IDTLKGVKSCHTGINTVGNVPVGYLVESGRSLVNGCDVLKAVSDYFGGSCVPGAGETS   | 180 |
| Db | 121 | IDTLKGVKSCHTGINTVGNVPVGYLVESGRSLVNGCDVLKAVSDYFGGSCVPGAGETS   | 180 |
| Qy | 181 | YSESRLCRGDSGEGVCDKSPLERYDYSGAFRLAEGAGDVAFKHSTVLENTDGT        | 240 |
| Db | 181 | YSESRLCRGDSGEGVCDKSPLERYDYSGAFRLAEGAGDVAFKHSTVLENTDGT        | 240 |
| Qy | 241 | LPSWGALLSQDFELLRCRGSRADVTWEQCHLARPAPAHAVVVRADTDGGLIFRLNNEG   | 300 |
| Db | 241 | LPSWGALLSQDFELLRCRGSRADVTWEQCHLARPAPAHAVVVRADTDGGLIFRLNNEG   | 300 |
| Qy | 301 | RLFSHSGSSFQMPSSSEAYGQKLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPN | 360 |
| Db | 301 | RLFSHSGSSFQMPSSSEAYGQKLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPN | 360 |
| Qy | 361 | RLPPYLRCVLTSTPEIKQCGMDMAVAFRRQLKPEIQCVSAKSPQHCMERIQAEQVDATL  | 420 |
| Db | 361 | RLPPYLRCVLTSTPEIKQCGMDMAVAFRRQLKPEIQCVSAKSPQHCMERIQAEQVDATL  | 420 |

```

QY 421 SGEDIYTAGKYGVLVPAAGEHAPEDSSNYVAVVVRDSSHAFTLDELGRKRSCHAGF 480
DB 421 SEDIYTAGKYGVLVPAAGEHAPEDSSNYVAVVVRDSSHAFTLDELGRKRSCHAGF 480
QY 481 GSPAGWDVPGALIQGFIRPKDCDVLTAVERSEFFNASCVPVNNPKNYPSSLCALCVGDEQ 540
DB 481 GSPAGWDVPGALIQGFIRPKDCDVLTAVERSEFFNASCVPVNNPKNYPSSLCALCVGDEQ 540
QY 541 GRNKCVCNSQERYGGRGAFRCFLVENAGDVAFVRHTTVPDNTNGHNSPWAELRSEDYE 600
DB 541 GRNKCVCNSQERYGGRGAFRCFLVENAGDVAFVRHTTVPDNTNGHNSPWAELRSEDYE 600
QY 601 LLCPCNGARAEVSQFAACNLAIQIPPHAVMVRPDTNIFTVYGLDKAODLFGDDHKNKGFKN 660
DB 601 LLCPCNGARAEVSQFAACNLAIQIPPHAVMVRPDTNIFTVYGLDKAODLFGDDHKNKGFKN 660
QY 661 FDSSNYHGODLLFKDQATVRAVPVGEKTYRGMGLDYVAALGEMSSQCSGAAAPAPGAP 720
DB 661 FDSSNYHGODLLFKDQATVRAVPVGEKTYRGMGLDYVAALGEMSSQCSGAAAPAPGAP 720
QY 721 LLPPLLPALAAARLLPPAL 738
DB 721 LLPPLLPALAAARLLPPAL 738

```

RESULT 2

```

US-08-520-933-3
; Sequence 3, Application US/08520933
; Patent No. 5981194
; GENERAL INFORMATION:
; APPLICANT: Jefferies, Wilfred A.
; APPLICANT: McGeer, Patrick L.
; APPLICANT: Rothenberger, Sylvia
; APPLICANT: Food, Michael R.
; APPLICANT: Yamada, Tatsuo
; APPLICANT: Kennard, Malcolm
; TITLE OF INVENTION: Use of p97 and Iron Binding Proteins
; TITLE OF INVENTION: as Diagnostic and Therapeutic Agents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Bereskin & Parr
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSH 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/520,933
; FILING DATE: August 31, 1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Shona S. McDiarmid
; REGISTRATION NUMBER: 38,798
; REFERENCE/DOCKET NUMBER: 7685-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-364-7311
; TELEFAX: 416-361-1398
; TELEX: 06-23115
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 719 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-520-933-3

```

Query Match 97.6%; Score 3840; DB 2; Length 719;

```

Best Local Similarity 100.0%; Pred. No. 0;
Matches 719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 GMEVWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVCQLIAAQADAITLDG 79
DB 1 GMEVWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVCQLIAAQADAITLDG 60
QY 80 GAIEAGKEHGLKPVVGEVYDQEVCTSYAVAVVRSSHVITDITLKVKSCHTGIRTVG 139
DB 61 GAIEAGKEHGLKPVVGEVYDQEVCTSYAVAVVRSSHVITDITLKVKSCHTGIRTVG 120
QY 140 MNVPVGYLVESGRSLVSMGCDVLKAVSDYFGSCVPGAGETSYSESICRLCRDSSGEGVC 199
DB 121 MNVPVGYLVESGRSLVSMGCDVLKAVSDYFGSCVPGAGETSYSESICRLCRDSSGEGVC 180
QY 200 DKSPLEYYDYGAFRCIAEAGDVAFVKHSTVLENTDGKTLPSWGQALLSQDFELLCRD 259
DB 181 DKSPLEYYDYGAFRCIAEAGDVAFVKHSTVLENTDGKTLPSWGQALLSQDFELLCRD 240
QY 260 GSRADVTEWROCHLARVPAHAVVVRADTDGGLIFRLINEGQRLFSHEGSSQFMSSEAYG 319
DB 241 GSRADVTEWROCHLARVPAHAVVVRADTDGGLIFRLINEGQRLFSHEGSSQFMSSEAYG 300
QY 320 QKDLLFKDSTSELVPIATQTYEAWLGHYELHAMKGLLCDPNRLPPYLRCVLSLSTPEIQK 379
DB 301 QKDLLFKDSTSELVPIATQTYEAWLGHYELHAMKGLLCDPNRLPPYLRCVLSLSTPEIQK 360
QY 380 GDMAVAFRRQRLKPEIQCVSAKSPQHMERIQAEQVDVATLSGEDIYTAGKYGVLVPAAG 439
DB 361 GDMAVAFRRQRLKPEIQCVSAKSPQHMERIQAEQVDVATLSGEDIYTAGKYGVLVPAAG 420
QY 440 EHYAPEDSSNYVAVVVRDSSHAFTLDELGRKRSCHAGFSPAGWDVPGALIQGF 499
DB 421 EHYAPEDSSNYVAVVVRDSSHAFTLDELGRKRSCHAGFSPAGWDVPGALIQGF 480
QY 500 RPKDCDVLTAVERSEFFNASCVPVNNPKNYPSSI.CALCVGDEOGRNKCVCNSQERYGYRGA 559
DB 481 RPKDCDVLTAVERSEFFNASCVPVNNPKNYPSSI.CALCVGDEOGRNKCVCNSQERYGYRGA 540
QY 560 FRCLVENAGDVAFVRHTTVPDNTNGHNSPWAELRSEDYELLCPCNGARAEVSPAAACNL 619
DB 541 FRCLVENAGDVAFVRHTTVPDNTNGHNSPWAELRSEDYELLCPCNGARAEVSPAAACNL 600
QY 620 AQIPPHAVMVRPDTNIFTVYGLDKAODLFGDDHKNKGFKNFDSNHYGODLLFKDQATVR 679
DB 601 AQIPPHAVMVRPDTNIFTVYGLDKAODLFGDDHKNKGFKNFDSNHYGODLLFKDQATVR 660
QY 680 AVPVGEKTYRGMGLDYVAALGEMSSQCSGAAAPAPCAPLLPLLPALAAARLLPPAL 738
DB 661 AVPVGEKTYRGMGLDYVAALGEMSSQCSGAAAPAPCAPLLPLLPALAAARLLPPAL 719

```

RESULT 3

```

US-09-285-040-3
; Sequence 3, Application US/09285040
; Patent No. 6455494
; GENERAL INFORMATION:
; APPLICANT: Jefferies, Wilfred A.
; APPLICANT: McGeer, Patrick L.
; APPLICANT: Rothenberger, Sylvia
; APPLICANT: Food, Michael R.
; APPLICANT: Yamada, Tatsuo
; APPLICANT: Kennard, Malcolm
; TITLE OF INVENTION: Use of p97 and Iron Binding Proteins
; TITLE OF INVENTION: as Diagnostic and Therapeutic Agents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSH 3Y2

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/285,040  
FILING DATE: 01-APR-1999  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gravelle, Micheline  
REGISTRATION NUMBER: 40,261  
REFERENCE/DOCKET NUMBER: 7685-032  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 416-364-7311  
TELEFAX: 416-361-1398  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-285-040-3

Query Match 97.6%; Score 3840; DB 4; Length 719;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 GMEVRCATSDPEQHKCGNMSEAFREAGIOPSLLCVVGTSADHCVOLIAAQEADAITLDG 79  
DB 1 GMEVRCATSDPEQHKCGNMSEAFREAGIOPSLLCVVGTSADHCVOLIAAQEADAITLDG 60  
QY 80 GAIYEAGKEHGLKPVVGEVYDQEVGTSYAVAVVRRSSHVTIDTLGKVSCHTGINRTVG 139  
DB 61 GAIYEAGKEHGLKPVVGEVYDQEVGTSYAVAVVRRSSHVTIDTLGKVSCHTGINRTVG 120  
QY 140 MNVPVGYLVESGRSLVWGCDVLKAVSDYFGGSCVPGAGETSYSESICRLCRGDSSEGVC 199  
DB 121 MNVPVGYLVESGRSLVWGCDVLKAVSDYFGGSCVPGAGETSYSESICRLCRGDSSEGVC 180  
QY 200 DKSPLERYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGTLPSPWQALLSQDFELLCRD 259  
DB 181 DKSPLERYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGTLPSPWQALLSQDFELLCRD 240  
QY 260 GSRADVTWRQCHLARVPAHAVVVRADTDGGLIFRLINEGORLFSHEGSSFQMFSEAYG 319  
DB 241 GSRADVTWRQCHLARVPAHAVVVRADTDGGLIFRLINEGORLFSHEGSSFQMFSEAYG 300  
QY 320 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAWKGLLCDPNRPPYLRCVLTSTPEIQKC 379  
DB 301 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAWKGLLCDPNRPPYLRCVLTSTPEIQKC 360  
QY 380 GMAVAFRRQRLKPEIQCVSAKSPQHCMERIOAEQVDVATLSEDIYTAGKYGVLVPAAG 439  
DB 361 GMAVAFRRQRLKPEIQCVSAKSPQHCMERIOAEQVDVATLSEDIYTAGKYGVLVPAAG 420  
QY 440 EHYAPEDSSNSYYVAVVRDSSHAFTLDELGRKRSCHAGFGSPAGWDVPVGALIORGFI 499  
DB 421 EHYAPEDSSNSYYVAVVRDSSHAFTLDELGRKRSCHAGFGSPAGWDVPVGALIORGFI 480  
QY 500 RPKDCDLVTAVSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKCVCNQSERYGYRGA 559  
DB 481 RPKDCDLVTAVSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKCVCNQSERYGYRGA 540  
QY 560 FRCLVENAGDVAFVRHTTVFDNTNGHNSPWAELRSEDEYELLCNPGARAEVSQFAACNL 619  
DB 541 FRCLVENAGDVAFVRHTTVFDNTNGHNSPWAELRSEDEYELLCNPGARAEVSQFAACNL 600  
QY 620 AQIPPHAVMVRPDNTNIFTVYGLLDKAQDLFGDDHKNKGKMFDSNHYGQDLLFKDATVR 679  
DB 601 AQIPPHAVMVRPDNTNIFTVYGLLDKAQDLFGDDHKNKGKMFDSNHYGQDLLFKDATVR 660  
QY 680 AVPVGEKTTYRGWGLDYVAALLEGSSQCSGAAAPAPGAPLPLLLPALAARLLPPAL 738

DB 661 AVPVGEKTTYRGWGLDYVAALLEGSSQCSGAAAPAPGAPLPLLLPALAARLLPPAL 719  
RESULT 4  
5262177-5  
Patent No. 5262177  
APPLICANT: BROWN, J OSEPH P.; ESTIN, CHARLES D.; PLOWMAN, GREGORY  
D. HELLSTROM, KARL E.; ROSE, TIMOTHY M.; HELLSTROM, INGERGERD;  
PURCHIO, ANTHONY F.; HU, SHIU-LOK; PENNATHUR, SRIDHAR  
TITLE OF INVENTION: RECOMBINANT VIRUSES ENCODING THE HUMAN  
MELANOMA-ASSOCIATED ANTIGEN  
NUMBER OF SEQUENCES: 6  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/7230  
FILING DATE: 27-JAN-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 827,313  
FILING DATE: 07-FEB-1986  
SEQ ID NO: 5  
LENGTH: 717  
5262177-5  
Query Match 96.3%; Score 3792; DB 6; Length 717;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 714; Conservative 1; Mismatches 2; Indels 2; Gaps 2;  
QY 20 GMEVRCATSDPEQHKCGNMSEAFREAGIOPSLLCVVGTSADHCVOLIAAQEADAITLDG 79  
DB 1 GMEVRCATSDPEQHKCGNMSEAFREAGIOPSLLCVVGTSADHCVOLIAAQEADAITLDG 60  
QY 80 GAIYEAGKEHGLKPVVGEVYDQEVGTSYAVAVVRRSSHVTIDTLGKVSCHTGINRTVG 139  
DB 61 GAIYEAGKEHGLKPVVGEVYDQEVGTSYAVAVVRRSSHVTIDTLGKVSCHTGINRTVG 120  
QY 140 MNVPVGYLVESGRSLVWGCDVLKAVSDYFGGSCVPGAGETSYSESICRLCRGDSSEGVC 199  
DB 121 MNVPVGYLVESGRSLVWGCDVLKAVSDYFGGSCVPGAGETSYSESICRLCRGDSSEGVC 180  
QY 200 DKSPLERYDYSGAFRCLEAGAGDVAFVKHSTVLENTDGTLPSPWQALLSQDFELLCRD 259  
DB 181 DKSPLERYDYSGA - RCLAEAGAGDVAFVKHSTVLENTDGTLPSPWQALLSQDFELLCRD 239  
QY 260 GSRADVTWRQCHLARVPAHAVVVRADTDGGLIFRLINEGORLFSHEGSSFQMFSEAYG 319  
DB 240 GSRADVTWRQCHLARVPAHAVVVRADTDGGLIFRLINEGORLFSHEGSSFQMF - FEAYG 298  
QY 320 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAWKGLLCDPNRPPYLRCVLTSTPEIQKC 379  
DB 299 QKDLLFKDSTSELVPIATQTYEAWLGHEYLHAWKGLLCDPNRPPYLRCVLTSTPEIQKC 358  
QY 380 GMAVAFRRQRLKPEIQCVSAKSPQHCMERIOAEQVDVATLSEDIYTAGKYGVLVPAAG 439  
DB 359 GMAVAFRRQRLKPEIQCVSAKSPQHCMERIOAEQVDVATLSEDIYTAGKYGVLVPAAG 418  
QY 440 EHYAPEDSSNSYYVAVVRDSSHAFTLDELGRKRSCHAGFGSPAGWDVPVGALIORGFI 499  
DB 419 EHYAPEDSSNSYYVAVVRDSSHAFTLDELGRKRSCHAGFGSPAGWDVPVGALIORGFI 478  
QY 500 RPKDCDLVTAVSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKCVCNQSERYGYRGA 559  
DB 479 RPKDCDLVTAVSEFFNASCVPVNNPKNYPSSLCALCVGDEQGRNKCVCNQSERYGYRGA 538  
QY 560 FRCLVENAGDVAFVRHTTVFDNTNGHNSPWAELRSEDEYELLCNPGARAEVSQFAACNL 619  
DB 539 FRCLVENAGDVAFVRHTTVFDNTNGHNSPWAELRSEDEYELLCNPGARAEVSQFAACNL 598  
QY 620 AQIPPHAVMVRPDNTNIFTVYGLLDKAQDLFGDDHKNKGKMFDSNHYGQDLLFKDATVR 679  
DB 599 AQIPPHAVMVRPDNTNIFTVYGLLDKAQDLFGDDHKNKGKMFDSNHYGQDLLFKDATVR 658  
QY 680 AVPVGEKTTYRGWGLDYVAALLEGSSQCSGAAAPAPGAPLPLLLPALAARLLPPAL 738

Db 659 AVPGKTTTGGWGLDYVAALGCMSSQCCSQAAPAPGAPLLPLLLPALAARLLPPAL 717

RESULT 5

US-08-459-818-19  
; Sequence 19, Application US/08459818  
; Patent No. 5851795

GENERAL INFORMATION:

APPLICANT: Linsley, Peter S.  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Damle, Nitin K.  
APPLICANT: Brady, William  
TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Merchant & Gould  
STREET: 11150 Santa Monica Blvd., Suite 400  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Fast-Seq 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,818  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Adriano, Sarah B.  
REGISTRATION NUMBER: 34,470  
REFERENCE/DOCKET NUMBER: 30436.35US02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 310-445-1140  
TELEFAX: 310-445-9031

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-459-818-19

Query Match 47.0%; Score 1850; DB 2; Length 502;

Best Local Similarity 100.0%; Pred. No. 3.1e-172;

Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |   |     |
|----|-----|---|-----|
| QY | 20  | GMEVWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG | 79  |
| Db | 156 | GMEVWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG | 215 |
| QY | 80  | GAIEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLKGKVSCHTGINRTVG   | 139 |
| Db | 216 | GAIEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLKGKVSCHTGINRTVG   | 275 |
| QY | 140 | WNVPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPGAGETSYESLCLRCRDSGEGVC   | 199 |
| Db | 276 | WNVPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPGAGETSYESLCLRCRDSGEGVC   | 335 |
| QY | 200 | DKSPLEYYDYGAFRCLEAGAGDVAFVKHSTVLENTDGTLPWSGQALLSQDFELLCRD   | 259 |
| Db | 336 | DKSPLEYYDYGAFRCLEAGAGDVAFVKHSTVLENTDGTLPWSGQALLSQDFELLCRD   | 395 |
| QY | 260 | GSRADVTETWROCHLARVPAHVVVRADTDGGLIFRLNEGQRLFSHGSSFQMFSSAYG   | 319 |
| Db | 396 | GSRADVTETWROCHLARVPAHVVVRADTDGGLIFRLNEGQRLFSHGSSFQMFSSAYG   | 455 |
| QY | 320 | QKDLFFKDSSTSELVPIATQTYEAWLGHYLAHMKGLLCDPNRPPYL 366          |     |
| Db | 456 | QKDLFFKDSSTSELVPIATQTYEAWLGHYLAHMKGLLCDPNRPPYL 502          |     |

RESULT 6

US-08-889-666-19  
; Sequence 19, Application US/08889666  
; Patent No. 5885579

GENERAL INFORMATION:

APPLICANT: Linsley, Peter S.  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Damle, Nitin K.  
APPLICANT: Brady, William  
APPLICANT: Kiener, Peter A.  
TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Merchant & Gould  
STREET: 11150 Santa Monica Blvd., Suite 400  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/889,666  
FILING DATE: 08-JUL-1997  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/375390  
FILING DATE: 18-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Adriano, Sarah B.  
REGISTRATION NUMBER: 34,470  
REFERENCE/DOCKET NUMBER: 30436-35US01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 310-445-1140  
TELEFAX: 310-445-9031

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-889-666-19

Query Match 47.0%; Score 1850; DB 2; Length 502;

Best Local Similarity 100.0%; Pred. No. 3.1e-172;

Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |   |     |
|----|-----|---|-----|
| QY | 20  | GMEVWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG | 79  |
| Db | 156 | GMEVWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG | 215 |
| QY | 80  | GAIEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLKGKVSCHTGINRTVG   | 139 |
| Db | 216 | GAIEAGKEHGLKPVVGEVYDQVGTSYAVAVVRRSSHVTIDTLKGKVSCHTGINRTVG   | 275 |
| QY | 140 | WNVPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPGAGETSYESLCLRCRDSGEGVC   | 199 |
| Db | 276 | WNVPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPGAGETSYESLCLRCRDSGEGVC   | 335 |
| QY | 200 | DKSPLEYYDYGAFRCLEAGAGDVAFVKHSTVLENTDGTLPWSGQALLSQDFELLCRD   | 259 |
| Db | 336 | DKSPLEYYDYGAFRCLEAGAGDVAFVKHSTVLENTDGTLPWSGQALLSQDFELLCRD   | 395 |
| QY | 260 | GSRADVTETWROCHLARVPAHVVVRADTDGGLIFRLNEGQRLFSHGSSFQMFSSAYG   | 319 |
| Db | 396 | GSRADVTETWROCHLARVPAHVVVRADTDGGLIFRLNEGQRLFSHGSSFQMFSSAYG   | 455 |

QY 320 QKDLFKDSTSELVPIATQTYEAWLGHEYLHAWKGLLCDPRLPPYL 366  
 Db 456 QKDLFKDSTSELVPIATQTYEAWLGHEYLHAWKGLLCDPRLPPYL 502

RESULT 7

US-08-465-078-19  
 ; Sequence 19, Application US/08465078  
 ; Patent No. 5885796  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Linsley, Peter S.  
 ; APPLICANT: Ledbetter, Jeffrey A.  
 ; APPLICANT: Damle, Nitin K.  
 ; APPLICANT: Brady, William  
 ; APPLICANT: Kiener, Peter A.  
 ; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
 ; NUMBER OF SEQUENCES: 26  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Merchant & Gould  
 ; STREET: 11150 Santa Monica Blvd., Suite 400  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 90025  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/465,078  
 ; FILING DATE: 05-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/375390  
 ; FILING DATE: 18-JAN-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Adriano, Sarah B.  
 ; REGISTRATION NUMBER: 34,470  
 ; REFERENCE/DOCKET NUMBER: 30436-35US01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 310-445-9031  
 ; TELEFAX: 310-445-1140  
 ; INFORMATION FOR SEQ ID NO: 19:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 502 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-465-078-19

Query Match 47.0%; Score 1850; DB 2; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-172;  
 Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 20 GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 79  
 Db 156 GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 215  
 QY 80 GAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDTLKGVSCHTGINTVVG 139  
 Db 216 GAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDTLKGVSCHTGINTVVG 275  
 QY 140 WNPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPCGAGETSYSESLCRLCRGDSGEGVC 199  
 Db 276 WNPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPCGAGETSYSESLCRLCRGDSGEGVC 335  
 QY 200 DKSPLEYYDYSGAFRLAEGAGDVAFAVKHSTVLENTDGTLPWSQALLSQDFELLCRD 259  
 Db 336 DKSPLEYYDYSGAFRLAEGAGDVAFAVKHSTVLENTDGTLPWSQALLSQDFELLCRD 395

QY 260 GSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQRLFSHEGSSFQMFSEAYG 319  
 Db 396 GSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQRLFSHEGSSFQMFSEAYG 455  
 QY 320 QKDLFKDSTSELVPIATQTYEAWLGHEYLHAWKGLLCDPRLPPYL 366  
 Db 456 QKDLFKDSTSELVPIATQTYEAWLGHEYLHAWKGLLCDPRLPPYL 502

RESULT 8

US-08-725-776-19  
 ; Sequence 19, Application US/08725776  
 ; Patent No. 5968510  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Linsley, Peter S.  
 ; APPLICANT: Ledbetter, Jeffrey A.  
 ; APPLICANT: Damle, Nitin K.  
 ; APPLICANT: Brady, William  
 ; APPLICANT: Kiener, Peter A.  
 ; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
 ; NUMBER OF SEQUENCES: 26  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Merchant & Gould  
 ; STREET: 11150 Santa Monica Blvd., Suite 400  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 90025  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/725,776  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/375390  
 ; FILING DATE: 18-JAN-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Adriano, Sarah B.  
 ; REGISTRATION NUMBER: 34,470  
 ; REFERENCE/DOCKET NUMBER: 30436-35US01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 310-445-1140  
 ; TELEFAX: 310-445-9031  
 ; INFORMATION FOR SEQ ID NO: 19:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 502 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-725-776-19

Query Match 47.0%; Score 1850; DB 2; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-172;  
 Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 20 GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 79  
 Db 156 GMEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQEADAITLDG 215  
 QY 80 GAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDTLKGVSCHTGINTVVG 139  
 Db 216 GAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVTIDTLKGVSCHTGINTVVG 275  
 QY 140 WNPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPCGAGETSYSESLCRLCRGDSGEGVC 199  
 Db 276 WNPVGYLVESGRSLVMGCDVLKAVSDYFGGSCVPCGAGETSYSESLCRLCRGDSGEGVC 335  
 QY 200 DKSPLEYYDYSGAFRLAEGAGDVAFAVKHSTVLENTDGTLPWSQALLSQDFELLCRD 259

Db 336 DKSPLEYYDYGAFRCLEAGAGDVAFVKHSTVLENTDGTLPWSGQALLSQDFELLCRD 395  
QY 260 GSRADVTEWROCHLARVPAHVVVRADTDGGLIFRLNEGQRLFSHEGSSFQMFSEAYG 319  
Db 396 GSRADVTEWROCHLARVPAHVVVRADTDGGLIFRLNEGQRLFSHEGSSFQMFSEAYG 455  
QY 320 QKDLFLKDSSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPRLPPYL 366  
Db 456 QKDLFLKDSSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPRLPPYL 502

## RESULT 9

US-08-488-062-19  
; Sequence 19, Application US/08488062  
; Patent No. 5977318  
; GENERAL INFORMATION:  
; APPLICANT: Linsley, Peter S.  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Damle, Nitin K.  
; APPLICANT: Brady, William  
; APPLICANT: Kiener, Peter A.  
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant & Gould  
; STREET: 11150 Santa Monica Blvd., Suite 400  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,062  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/375390  
; FILING DATE: 18-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adriano, Sarah B.  
; REGISTRATION NUMBER: 34,470  
; REFERENCE/DOCKET NUMBER: 30436-35US01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 310-445-1140  
; TELEFAX: 310-445-9031  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 502 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-488-062-19

Query Match 47.0%; Score 1850; DB 2; Length 502;  
Best Local Similarity 100.0%; Pred. No. 3.1e-172;  
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 20 GNEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQADAITLDG 79  
Db 156 GNEVRCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSADHCVQLIAAQADAITLDG 215  
QY 80 GAIYEAGKHGLKPVVGVYDQEVGYSYAVAVVRSSHVTIDTLKGVKSCHTGNRTVG 139  
Db 216 GAIYEAGKHGLKPVVGVYDQEVGYSYAVAVVRSSHVTIDTLKGVKSCHTGNRTVG 275  
QY 140 WNPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYSESICRLCRGSSGEGVC 199

Db 276 WNPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETSYSESICRLCRGSSGEGVC 335  
QY 200 DKSPLEYYDYGAFRCLEAGAGDVAFVKHSTVLENTDGTLPWSGQALLSQDFELLCRD 259  
Db 336 DKSPLEYYDYGAFRCLEAGAGDVAFVKHSTVLENTDGTLPWSGQALLSQDFELLCRD 395  
QY 260 GSRADVTEWROCHLARVPAHVVVRADTDGGLIFRLNEGQRLFSHEGSSFQMFSEAYG 319  
Db 396 GSRADVTEWROCHLARVPAHVVVRADTDGGLIFRLNEGQRLFSHEGSSFQMFSEAYG 455  
QY 320 QKDLFLKDSSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPRLPPYL 366  
Db 456 QKDLFLKDSSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPRLPPYL 502

## RESULT 10

US-08-175-158A-2  
; Sequence 2, Application US/08175158A  
; Patent No. 5986067  
; GENERAL INFORMATION:  
; APPLICANT: FUNK, Walter D.  
; APPLICANT: MacGILLIVRAY, Ross T.A.  
; APPLICANT: MASON, Anne B.  
; APPLICANT: WOODWORTH, Robert C.  
; TITLE OF INVENTION: RECOMBINANT TRANSFERRINS, TRANSFERRIN HALF-  
; TITLE OF INVENTION: MOLECULES AND MUTANTS THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/175,158A  
; FILING DATE: 28-DEC-1993  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/832,029  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A.  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/DOCKET NUMBER: UVI-005CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 698 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-175-158A-2

Query Match 32.3%; Score 1269.5; DB 2; Length 698;  
Best Local Similarity 40.4%; Pred. No. 3.9e-115;  
Matches 304; Conservative 118; Mismatches 221; Indels 109; Gaps 32;  
QY 1 MRGPGSALWL--LLALRTVLGMEYRWCATSDPEQHKCGNMSEAFRE-----AGIQPSL 52  
Db 1 MRLAVGALLVCAVLGLCLAVPKTVRCVAVSEHEATKC---QSPRDMKSVIPSDGPSV 56  
QY 53 LCVRTGTSADHCVQLIAAQADAITLDGGAIVYAG--KEHGLKPVVCEVY--DOEVTSYYA 109  
Db 57 ACVKKASYLDICITRAIANNEADAVTLDAVLVDAYLAPNNLKPVVAEFGYSGKEDPQTYYA 116

|    |     |  |
|----|-----|--|
| Qy | 110 | VAVRRSSHVTTDTLKGUKSCHTGNRTVGMNVPVGVILNBSGRLSVMGCD-----VLK 162      |
| Db | 117 | VAVVKDGSFQMNQLRGKSGHITGLGRSAGNWIPIGLY-----CDLPEPRPLEK 167          |
| Qy | 163 | AVSDYFGGSCVPGAGETSYSESICRLCRGSDSGEGVCDKSPLERYDYSGAFRCLEAGAG 222    |
| Db | 168 | AVANFFSGSCAPCADGTDFPQ-LCQLCFG-----CGCSTLNQYFGYSARFKCLDKGAG 219     |
| Qy | 223 | DVAFVVKHSTVLENTDGTGLPFGSQALLSQD-FELLCRDGSRADVTYEQRCHLARVPAHAV 281  |
| Db | 220 | DVAFVVKHSTIFENLANKA-----DRDQVELLCLDNTRKPVDEYKDCHLAQVPSHTV 270      |
| Qy | 282 | VVRADTDGG--LIIFRLNNEGORLPSHEGS-SFOWESSEAGOKDLLFKDSTSELVPIAT 337    |
| Db | 271 | VAR--SMGKEDLIWELINQAQHEGKDKSEFQLFSS-PHG-KOLLFKDQSAHGFLKRVPP 326    |
| Qy | 338 | Q-TYEAWLGHYELHAKMGLL-----CDPNRLPPYLRLWCVLSTPBIQRKCGDMAVA 385       |
| Db | 327 | RMDAKMYLGYEYVTAIRNLREGTCPEAPTDECKP-----VKWCALSHERLAKCDEWSV- 379    |
| Qy | 386 | FRQRRLKPIQCVSAKSPOHCHMERLOARQVDAVTLSGEDIYTAGKYGVLPAAGEHYAPE 445    |
| Db | 380 | -----NSVGKIECVSAETTEDCIATKMGNEADAMLDGGFVYIAG-KOGLVPFLVAENYKS 434   |
| Qy | 446 | DGSN-----SYVJVAVRRDSDHAFTDLRGKSKSHAGFGSPAGWDVPVGCALIQRGFKR 500     |
| Db | 435 | DNCEDTPEAGYFAVAVVKXSAD-ITWNLLGKKSKCHTAVGRTAGWNIPIGLI-----YNK 489   |
| Qy | 501 | PKDCDVLTAVSFFFNASCVPVNNPKNYPSSICALCVGDBOGRNKCVCNGSQERYGYRGAF 560   |
| Db | 490 | INHC-----RFDEFFSEGCAPSKK---DSSLCKLMG-SGLNLCEPNKKEGYGYGTGAF 540     |
| Qy | 561 | RCLVENAGDVAFVRHTTVPDNTNGHNSPWAELRSEDYELLCPNGARASVSGFAACNLA 620     |
| Db | 541 | RCLVEK-GDVAFVKHQTPQNTGGKKNPDPWAKNLNEXDYELLCLDGRKFEVEEYANCHLA 599   |
| Qy | 621 | QTPPHAVMVRPDTNITFTYVGLLDKADQLFDGDDHN--KNGFKMFDSSNYHGDQLLFKDQTV 678 |
| Db | 600 | RAPNAHVTRDKDEA-CVHKILRQOHLFGSNVTDGSGNCFLFSFET---KDLLFRDDTV 655     |
| Qy | 679 | RAVPVGEKTTYGWLGLDYVAALGEMSSQQCS 710                                |
| Db | 656 | CLAKLHDRNTYEKYLGEYVRAVGML--RKGS 685                                |

## RESULT 11

US-08-470-058-2

Sequence 2, Application US/08470058

Patent No. 5817789

GENERAL INFORMATION:

APPLICANT: Heartlein, Michael W.

APPLICANT: Lemont, Jeffrey F.

TITLE OF INVENTION: Chimeric Proteins For Use in Transport

TITLE OF INVENTION: of a Selected Substance Into Cells

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,058

FILING DATE: 06-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

RESULT 12

US-09-037-188-2





QY 645 AQDLFGDDHN--KNGFKMFDSSNYHGQDLFLFKDATVRAVPVGEKTTYRGWLGLDYVAALE 702  
 Db 1335 QQHFLGNSVNTDCSGNFCFLRSET---KDLFRDDTVCLAKLHNRNTYKYLGEYVKAAG 1391  
 QY 703 GMSQQCS 710  
 Db 1392 NL--RKCS 1397

RESULT 15

US-09-037-188-4  
 ; Sequence 4, Application US/09037188  
 ; Patent No. 6027921  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Heartlein, Michael W.  
 ; APPLICANT: Lemontt, Jeffrey F.  
 ; APPLICANT: Concino, Michael F.  
 ; TITLE OF INVENTION: CHIMERIC PROTEINS FOR USE  
 ; TITLE OF INVENTION: IN TRANSPORT OF A SELECTED SUBSTANCE INTO CELLS  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: Windows95  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/037,188  
 ; FILING DATE: 02-MAR-1998  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE: 06-JUN-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fraser, Ph.D., J.D., Janis K.  
 ; REGISTRATION NUMBER: 34,819  
 ; REFERENCE/DOCKET NUMBER: 07236/009002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617/542-5070  
 ; TELEFAX: 617/542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1410 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-09-037-188-4

Query Match 32.1%; Score 1263.5; DB 3; Length 1410;  
 Best Local Similarity 40.8%; Pred. No. 4.9e-114;  
 Matches 297; Conservative 115; Mismatches 209; Indels 107; Gaps 31;  
 QY 23 VRWCATSDPEQHKCGNMSAFRE-----AGIQPSLLCYRGTSADHCVOLIAAQEADAIT 76  
 Db 737 VRWCATSEHATKC-----QSFDRHMKSVIPSDGPSVACVKKASYLDCIRAIANEADAVT 792  
 QY 77 LDGAIYEAG-KEHGLKPVVGEVY--DQEVGTSYAVAVVRRSSHVTIDTLKGVKSCHTG 133  
 Db 793 LDAGLVYDAYLAPNNLKPVVABFYGSKDPQTFYAVAVVKDSDGFMNQLRGKSCHTG 852  
 QY 134 INRTVGMNVPGVYLVESGRSLVWGC-----VLKAVSDYFGGSCVPGAGETSYSESIC 186  
 Db 853 LGRSAGWNPIGLLY-----CDLPEPRKPLEKAVANFFSGSCAPCADGTDFFQ-JC 902  
 QY 187 RLCRGDSGEGVCDKPLERYDYSGAFCLAEAGADVAFVXHSVTLENTDKTLPSWQ 246  
 Db 903 QLCPG-----CGCSTLNQYFGYSGAFKCLKDAGADVAFVXHSVTIFENLANKA----- 949

Search completed: May 14, 2004, 09:44:34  
 Job time : 26 secs

QY 247 ALLSOD-FELLCRGSRADVTETWROCHLARPAPAHAVVVRADTDGG---LIFRLNEGORL 302  
 Db 950 ---DRDQYELLCLDNTRKPVDEYKDWCHLAQVPSHTVVAR--SMGSKEDLIWELNQAQEH 1004  
 QY 303 FSHEGS-SFQMFSEAYGQKDLLPKDSTSELVPIATQ-TYEAWLGHGHEYLHAKMGLL---- 356  
 Db 1005 FGKDKSKEFOLFSS-PHG-KDLLPKDSAHGFLKVPPRMDAKWYLGVEYVTAIRNLREGTC 1062  
 QY 357 -----CDPNRLPPYLRCVLSLSTPEIOKCGDMAVAFRRQRILKPEIOCVSAKSPHOCMER 409  
 Db 1063 PEAPTDECKP-----VKWCALSHHERLKCDWSV-----NSVGKIECVSAETTEDCIAK 1111  
 QY 410 IQAEQVDAVTLGSDIYTAGKKYGLVPAAGBHYAPEDSSN-----SYVVAVVRDSSHA 464  
 Db 1112 IMNGEADANSLDGGFVYIAG-KCGLVPVLAENYKNSDNCEDTPEAGYFAVAVVKASAD- 1169  
 QY 465 FTLDLGRKRSCHAGFGSPAGWDVPVGALIQGFIRPKDCDVLTAVERSEFFNASCVPVNNP 524  
 Db 1170 LTWDNLKGGKSCHTAVGRTAGWNIPMGLL-----YMKINHC-----RFDEFFSEGCAPGSKK 1221  
 QY 525 KNPSSSLCALCVGDDEQGRNKCNGSQERYGYRGAFRCCLVENAGDVAFVRHTTTFDNTNG 584  
 Db 1222 ---DSSLCKLQWG--SGLNLCEPNNKEGYGYGTGAFRCCLVEK-GDVAFAVKGHTVPQNTGG 1275  
 QY 585 HNSEPWAAELRSEDYELLCPNGARAEVSQFAACNLAIQIPPHAVMVRPDTNIFTVYGLLDK 644  
 Db 1276 KNPDPAKKNLNEKYELLCLDGTGTRKPVVEYANCHLAPNHAHVTRKDKEA-CVHKILRQ 1334  
 QY 645 AQDLFGDDHN--KNGFKMFDSSNYHGQDLFLFKDATVRAVPVGEKTTYRGWLGLDYVAALE 702  
 Db 1335 QQHFLGNSVNTDCSGNFCFLRSET---KDLFRDDTVCLAKLHNRNTYKYLGEYVKAAG 1391  
 QY 703 GMSQQCS 710  
 Db 1392 NL--RKCS 1397